Giancarlo Macchi Jánica

Spazio e Misura

Introduzione ai metodi geografico-quantitativi applicati allo studio dei fenomeni sociali

unisi Manuali

unisi Manuali

1

Giancarlo Macchi Jánica

Spazio e misura

introduzione ai metodi geografico-quantitativi
applicati allo studio dei fenomeni sociali

Si ringraziano: il Lap&t per l'autorizzazione all'uso delle fotografie oblique riprodotte alle figure 1.1 (Castello di Murlo) e 1.4 (Valle dell'Albegna); l'*American Geographical Society* per l'autorizzazione all'uso della figura 5 a pagina 244 del saggio di SVIATLOVSKY, EELLS, «Geographical Review», vol. 27, n. 2, 1937; la Taylor & Francis e la *Association of American Geographers* per l'autorizzazione all'uso della figura 3 a pagina 314 del saggio di HORVATH, «Annals of the Association of American Geographers», vol. 59, n. 2, 1969 e della figura 7 a pagina 234 del saggio di RODGERS, «Annals of the Association of American Geographers», vol. 64, n. 2, 1974 .

Titolo: *Spazio e misura: introduzione ai metodi geografico-quantitativi applicati allo studio dei fenomeni sociali*
Autore: Giancarlo Macchi Jánica

©Università degli Studi di Siena 2009

Edizioni dell'Università
ISBN: 978-88-96151-01-3

Grafica di copertina a cura dell'Ufficio Comunicazione on-line

Indice

Premessa **9**

Introduzione **13**
Obiettivi 14
I "problemi" di ricerca 17
L'interdisciplinarietà 18
I numeri e i linguaggi formali 19
Il calcolatore e il *software* 22

1 Ricerca storica e geografia quantitativa **25**
1.1 Introduzione all'analisi spaziale 25
1.2 L'analisi spaziale e la ricerca scientifica 28
1.3 La distribuzione spaziale 29
1.4 L'importanza della posizione nella ricerca sull'uomo 32
1.5 La complessità spaziale 36
1.6 Lo scopo dell'analisi spaziale 42
1.7 Analisi spaziale e ricerca storica 50

2 Modelli geografici **55**
2.1 La natura dei modelli scientifici 56
2.1.1 La costruzione dei modelli 59
2.2 Modelli geografici 65
2.3 Carte come modelli 67
2.4 Modello dei dati e cartografia digitale 69
2.4.1 Il modello vettoriale 69
2.4.2 Il modello *raster* 71

3 Analisi locazionale **73**
3.1 I poligoni di Thiessen o di Voronoy . . . 78
3.1.1 «Lo stato isolato» di Von Thünen . . . 84
3.2 La teoria della localizzazione delle industrie . . . 92
3.3 Il modello delle località centrali . . . 102
3.4 Il modello di Lösch . . . 109
3.5 La regola rango dimensione . . . 114

4 Statistica descrittiva in geografia **125**
4.1 Le misurazioni . . . 125
4.1.1 La misurazione spaziale . . . 127
4.2 Le scale di misurazione . . . 127
4.2.1 La scala nominale . . . 129
4.2.2 La scala ordinale . . . 132
4.2.3 Scala ad intervalli . . . 133
4.2.4 Scala a rapporti . . . 134
4.3 Le variabili . . . 135
4.3.1 Variabili categoriali o mutabili . . . 136
4.3.2 Variabili numeriche e quantitative . . . 136
4.3.3 Notazione delle variabili . . . 136
4.4 Il campionamento . . . 137
4.4.1 Campionamento della popolazione . . . 138
4.4.2 Campionamento del territorio . . . 139
4.5 Statistica descrittiva . . . 141
4.6 La classificazione dei dati e le distribuzioni di frequenze . . . 142
4.6.1 La costruzione delle classi . . . 146
4.6.2 Confini e valore centrale delle classi . . . 148
4.6.3 Istogrammi di frequenza . . . 149
4.6.4 Classificazione cartografica . . . 152
4.6.5 Curve cumulative . . . 154
4.7 Misure di "tendenza centrale" . . . 155
4.7.1 La media . . . 156
4.7.2 La mediana . . . 159
4.7.3 La moda . . . 161
4.7.4 Il *midrange* . . . 162
4.8 Misure di dispersione . . . 162
4.8.1 Campo di variazione . . . 164
4.8.2 I quartili . . . 164
4.8.3 La media interquartile . . . 166
4.8.4 La varianza . . . 166
4.8.5 Lo scarto quadratico medio . . . 168

4.9 Le distribuzioni statistiche 170
4.9.1 Probabilità e insediamento 171
4.9.2 La distribuzione normale 173
4.9.3 La distribuzione di Poisson 177

5 La densità **183**
5.1 Calcolo della densità 184
5.2 L'uso integrato della densità 188
5.3 Campioni geografici 190
5.4 Il test del chi–quadrato (χ^2) 192
5.5 Il χ^2 e la geografia quantitativa 197

6 Point pattern analysis **203**
6.1 Le carte per punti 204
6.1.1 Le carte a punti e la coppia di variabili x y 206
6.2 Il baricentro 206
6.2.1 Deviazione della distanza standard 208
6.2.2 L'utilizzo del baricentro nella ricerca storica 210
6.3 La dispersione spaziale 212
6.3.1 Dispersione spaziale e scala cartografica 216
6.4 Analisi a quadrati 217
6.4.1 Limiti dell'analisi a quadrati 222
6.5 Il vicino più prossimo 224
6.5.1 Test di significatività di R 231
6.5.2 Limiti del vicino più prossimo 233
6.6 La funzione K di Ripley 235
6.6.1 Proprietà di secondo ordine 235
6.6.2 Lo studio di *patterns* rurali con la funzione K di Ripley 241

7 Riflessioni conclusive **245**
7.1 Ostacoli sul percorso 246
7.2 I *patterns* 248
7.3 GIS e geografia quantitativa 248
7.4 La complessità 250

Riferimenti bibliografici **253**

Elenco delle figure **269**

Premessa

Spazio e misura: si può ben dire che fin dalle origini è questo il tema centrale della geografia. La misura delle coordinate terrestri, da Eratostene in poi, ha consentito di costruire la carte dell'ecumene e poi un po' alla volta delle sue regioni. Nel XVIII secolo si passa, secondo una definizione di Alexandre Koyré, «dal mondo del pressappoco all'universo della precisione». La carta diviene solo allora uno strumento scientifico di rappresentazione dello spazio. E già allora si pose il problema di cui questo libro si occupa: e poi? Quali sono i contenuti che si vogliono rappresentare nello spazio definito della carta? Di qui la storia delle osservazioni e delle relative misurazioni, che richiedevano non solo strumenti sempre più affidabili ma anche la capacità di selezionare ciò che va osservato, per poi procedere alla combinazione delle misure in un linguaggio coerente. Il ruolo della geografia non è mai stato quello di spiegare i fenomeni, ma quello di registrarne la distribuzione spaziale in modo tale da suggerire domande e offrire opportunità di verifica. Osservare, misurare, combinare: questo il metodo del geografo fin dai tempi di Humboldt, quando la disciplina non aveva ancora pretese accademiche e poteva dialogare con la storia naturale e con la storia *tout-court*. Quando i dati erano pochi, e andavano raccolti con attenzione o anche – spesso e volentieri – inventati o almeno dedotti sulla base del linguaggio cartografico. E oggi? Non sarà che l'immensa mole di dati a disposizione ci ha fatto ritornare al mondo del pressappoco?

Questo libro colma un vuoto, o meglio un diaframma, che continua ancora oggi a tenere lontane le discipline storiche e sociali dai metodi quantitativi. Ancora oggi, cioè quando con gli strumenti dell'informatica questi metodi sono alla portata di tutti, giovani e meno giovani.

Per chi lavora in una facoltà di Lettere, in particolare, è facile constatare che il percorso di apprendimento legato alla comprensione della teoria e dei metodi quantitativi nell'ambito della geografia e delle scienze del territorio è particolarmente tortuoso e affidato spesso all'impegno individuale dei singoli, senza alcun appoggio concreto nell'ordinamento degli studi. A queste difficoltà si lega lo scarso interesse generale da parte della comunità delle scienze umane all'applicazione di un approccio formale nell'esposizione e nell'analisi dei

dati. Si preferiscono le derive culturaliste, facili scappatoie per raggiungere rapidamente brillanti risultati sul piano narrativo, che tuttavia lasciano spesso il tempo che trovano. I naturalisti citano in proposito le *Just so stories* di Rudyard Kipling, le "storie proprio così", dove si spiegava l'origine della proboscide dell'elefante o del collo della giraffa ricorrendo, appunto, al linguaggio narrativo: brillante quanto indimostrabile.

Le difficoltà non vanno cercate in un'assenza reale di manuali e testi: anche se esistono guide in lingua italiana (nella maggior parte dei casi tradotte dall'inglese), si tratta di una babele composta da modelli, funzioni ed equazioni, espresse con notazioni e linguaggi dissimili; fattore che irrimediabilmente finisce per aggravare una situazione già difficile di per sé: che le pagine di questo libro, almeno questo è il modesto auspicio dell'autore, non dovrebbero contribuire a peggiorare.

Nell'esporre gli elementi formali quali la notazione estesa di formule o equazioni, Macchi ha cercato infatti di usare un lessico elementare e chiaro, alternando al discorso numerico il linguaggio naturale. Nel fare ciò avrà forse saltato molti passaggi (soprattutto di ambito statistico), certo importanti ma trascurabili in una fase introduttiva ai metodi quantitativi. Ma anche se certe affermazioni – a livello formale – appariranno a qualcuno semplicistiche o magari imprecise, il testo unisce con successo lo sforzo di mantenere il rigore e la coerenza necessari ad un approccio coerente, con quello di rendere la lettura il più facile e semplice possibile. Si spera che il lettore abbia modo di colmare in un secondo momento eventuali lacune grazie ai riferimenti bibliografici.

Personalmente posso dire di aver trovato nell'esposizione di Macchi proprio il giusto ordine con il quale deve essere organizzato il percorso fra la grande mole di dati e fonti che oggi sono disponibili ma che rischiano di sfuggire di mano. Se non si possiede, ad esempio, una familiarità con parti della statistica descrittiva, la comprensione di certe procedure di analisi spaziale risultano enormemente difficoltose; e ciò comporta, come spesso succede, un abbandono immediato di qualsiasi intento.

D'altra parte, leggendo con la dovuta attenzione queste pagine, ci si rende conto che la metodologia non può essere considerata come un corpo statico e fisso. Essa presenta una sua evoluzione soggettiva nella maggior parte dei casi.

Nella stesura di questo volume Macchi ha voluto illustrare la propria visione personale sulla questione dei metodi della geografia quantitativa riorganizzando e integrando le idee ed i concetti nell'ordine ideale secondo il quale, con il senno di poi, lui stesso avrebbe voluto acquisirli: a distanza di cinquant'anni, devo ammettere che anch'io avrei voluto poter contare, quando mi sono formato allo studio del territorio, su una guida di questo genere. In altre parole, guardando indietro, ecco la guida che, come sarebbe stato utile a me, potrebbe

orientare gli studenti a muoversi nel labirinto delle misure: per acquisire un linguaggio appropriato e forse anche risparmiare tempo. Non si tratta di un volume totale delle cose importanti per lo studioso dei paesaggi storici. È solo una parte: quella introduttiva ed essenziale per avviare nel modo più facile un proprio percorso all'interno dell'analisi spaziale.

Si tratta quindi di un contributo a tutti coloro che, nel desiderio di comprendere la natura, le forme e l'evoluzione storica dell'occupazione sociale dello spazio, si accingono a costruire il proprio percorso di conoscenza o a sperimentarne dei nuovi.

Claudio Greppi

Introduzione

Il tentativo di scrivere un manuale di geografia quantitativa per le scienze umane non va confuso con quello di realizzare un compendio di "come le cose devono o dovrebbero essere fatte". Sarebbe un errore considerare un corpo metodologico come una sorta di valigetta degli attrezzi da cui estrarre utensili adatti a questa o quella indagine. È più opportuno considerare invece il presente testo come un insieme di note pratiche relative alle applicazioni di metodi geografici analitico-quantitativi[1] basate sull'esperienza dell'autore.

Pensai a questo testo per la prima volta cinque anni fa, durante il mio primo anno di dottorato di ricerca a Siena[2], ma altre attività mi distolsero da quel progetto, a cui non mi dedicai più fino ad ora. Nonostante il lungo intervallo di tempo, l'impalcatura del lavoro rispecchia ancora l'idea originaria: un testo in cui l'analisi locazionale e l'analisi spaziale potessero fondersi in un'unica guida, utile a studenti e studiosi interessati all'applicazione dei metodi quantitativi nell'indagine dei paesaggi umani e dei paesaggi storici con un approccio prevalentemente su scala regionale. Il risultato finale, ad eccezione di qualche variazione o integrazione, resta fedele al progetto iniziale, sul quale nel 2001 cominciai a scrivere alcune di queste righe. Parte di questo primo tentativo fu l'articolo *Introduzione all'analisi spaziale*, pubblicato nel primo quaderno «Trame nello Spazio»[3].

Il titolo del testo fu definito solo nel 2007, quando ho ripreso a lavorare alla realizzazione del manuale. Ad una prima analisi il titolo può apparire assai poco concreto, ma benché non si faccia riferimento all'uomo è lui – o meglio l'insediamento umano – al centro del discorso; questa idea di fondo

[1] Per una definizione dei "metodi analitico-quantitativi" in geografia consultare Vagaggini, Dematteis 1976, pp. 1-2.

[2] Macchi 2004. Gran parte dell'impalcatura espositiva proviene da questo elaborato, modificato più volte sotto forma di dispense e sottoposto ad un attento riesame nella presente veste. Parte dei risultati della tesi sono stati già pubblicati in Macchi 2007. Ringrazio Claudio Greppi, Antonio Moro e Anna Guarducci per il loro prezioso aiuto.

[3] Macchi 2003. Questo breve saggio, che doveva costituire l'introduzione del progetto originale, viene riproposto in questa sede rivisto e aggiornato.

viene esplicitata nel sottotitolo. Se l'astrattezza del titolo può sembrare poco opportuna, soprattutto per una guida universitaria, esso, come verrà chiarito lungo l'intero testo, mi sembra che riesca a ben sintetizzare il senso stesso di quello che doveva essere lo scopo già del progetto originario: l'attenzione va e deve essere rivolta allo spazio, tema spesso dimenticato dalle discipline umane. Ma misurare lo spazio non significa giungere a delle valutazioni assolute. La statistica e la matematica non costituiscono necessariamente sinonimo di precisione o obiettività e un approccio analitico non deve costituire l'obiettivo primario di per sé. Quel che ci si augura, facendovi ricorso, è invece di potere procedere, a prescindere dal metodo, alla formulazione di ipotesi che, in un modo o nell'altro, possano essere riprese, lette, valutate o confutate dalla comunità scientifica nel suo complesso.

Il concetto di "misura" inserito nel titolo va letto come un invito ad una dialettica più ampia e concreta su un elemento così importante come lo spazio. Suggerire l'utilizzo del linguaggio formale (cioè quello matematico) risponde alla necessità di un dialogo[4] reale e pratico sul tema delle forme d'occupazione sociale dello spazio. I modi e i linguaggi, per l'appunto, sono solo secondari; anche se è innegabile che il loro utilizzo da parte di una determinata comunità suggerisce anche il livello di consapevolezza che questa possiede del suo oggetto di studio, così come il grado di evoluzione di problematiche e domande scientifiche che su questo o quell'argomento è giunta a formulare.

Il soggettivismo e il relativismo sono forse le cause per le quali le discipline dell'uomo tendono a procedere in modo così lento. Non si illuda il lettore: anche l'adozione di un approccio formale può essere anch'esso molto soggettivo; in questo caso il vantaggio sta però nel fatto che, nel processo di comunicazione, la comunità riesce immediatamente nell'identificazione del grado o livello di relativismo implicito[5].

Obiettivi

In passato, nel fornire spiegazioni o indicazioni a chi (appartenente a settori disciplinari quali la storia, storia degli insediamenti, archeologia dei paesaggi, ecologia storica e geografia) si avvicinasse ai metodi quantitativi di tipo spaziale, ho avvertito in modo netto i limiti delle risposte. Nel precisare sempre

[4] Qui inteso come uno scambio di idee da parte di una comunità scientifica.

[5] Sul perché dei linguaggi formali nel campo della geografia si rimanda a Silk 1979, pp. 3-5: «Statistical methods are used by geographers because they help us to come to conclusions based upon empirical data, these being measurements derived either from observation or experiment. Geography was a strongly empirical subject well before the "quantitative revolution" of the 1960s, and so statistical methods can be regarded as essential aids to geographic enquiry.»

come i *Geographic Information Systems* (GIS) ricoprano per ora prevalentemente un ruolo strumentale, di fatto non molto diversi da una matita o una calcolatrice, mi addentravo in questioni relative all'applicazione dei metodi di quantificazione spaziale.

D'altronde la rivoluzione dei GIS ha fatto sì che, nel corso degli ultimi vent'anni, diverse banche dati di tipo spaziale siano state effettivamente sviluppate nell'ambito di diverse discipline. Fa parte della naturale evoluzione della ricerca, dunque, giungere alla fase di analisi dei dati: una volta ultimata la loro raccolta e archiviazione, dovrebbe infatti sorgere spontaneo domandarsi cosa poter fare di quelle informazioni.

Il processo d'analisi non è immediato o automatico. Al di là dei luoghi comuni, le banche dati non suggeriscono domande, nè tanto meno i modi migliori di porle[6].

Sentire il limite dell'esposizione strumentale significa che certe indicazioni, pur effettive nel risolvere problemi (per così dire, tecnici) non funzionavano o comunque non rilevavano gli aspetti fondamentali della ricerca scientifica di fondo. La portata della questione pratica non va oltre il superamento stesso di uno scoglio nel percorso dell'applicazione strumentale.

Lo scopo del libro, per quanto possibile, sarà quello di riorganizzare, in un unico corpo coerente dedicato alle discipline storiche e sociali, una serie eterogenea di nozioni metodologiche sparpagliate in più settori della geografia umana e della geografia economica. Ecco la difficoltà maggiore nel realizzare questo testo: si tratta di un intreccio che cerca, almeno nei propositi, di mantenere un collegamento con quelle domande centrali e le problematiche scientifiche di cui si diceva sopra; per prima la comprensione dello spazio umano attraverso una sua lettura semplificata. Questo sforzo viene fatto nell'intento di fondere e rendere coerenti tecniche e concetti ricavati da settori quali cartografia digitale, cartografia tradizionale, statistica, geometria, analisi spaziale, geostatistica, informatica, per citare solo le principali.

Il volume è stato scritto pensando a studenti universitari di diversi corsi di laurea, in particolare a quelli dei corsi di geografia, storia, lettere, scienze dei beni archeologici e culturali e architettura. Questo perché i settori disciplinari dell'area umana e sociale hanno un vincolo con lo spazio e gli aspetti ad esso legati: come un certo fenomeno si distribuisce nello spazio, oppure quale sia l'effetto dello spazio nell'articolazione del fenomeno stesso, ecc. Prova del legame di discipline quali storia, geografia umana, antropologia sociale, antro-

[6] Popper 1995b, p. 208: «Ricordo un'occasione in cui uno scienziato famoso presentò tale concezione – io ero presidente del *meeting*. La scienza, egli disse, consiste semplicemente nel misurare e correlare i risultati. Nella discussione che seguì, suggerii di richiedere al British Museum il permesso di misurare la lunghezza, larghezza, spessore e peso dei libri lì conservati, e di studiare quindi le possibili relazioni tra le misure ottenute.»

pologia culturale, sociologia e archeologia con gli aspetti spaziali, è il fatto che ognuna di queste si occupa, fra l'altro, anche della produzione di materiale cartografico. Questo rapporto può essere passivo o attivo, inconscio o conscio, ma sono veramente pochi i settori disciplinari che possono ignorare lo spazio.

Lo spazio costituisce da sempre uno dei pilastri fondamentali del pensiero occidentale. "Posizione", "distanza", "orientamento" o "forma" sono concetti basilari di qualsiasi ragionamento scientifico sulla distribuzione dei fenomeni. Sarà nel XVIII secolo lo stesso Kant ad elevare lo spazio al livello di categoria[7].

L'idea di fondo è che, in sostanza, le discipline dell'uomo debbano in qualche misura occuparsi del suo spazio. Semplicemente capire le logiche distributive dei gruppi sociali[8], così come semplicemente tutte le scienze applicate al regno minerale, vegetale e animale si sono impegnate nella comprensione dei *patterns* a livello spaziale. Non vi è motivo per pensare che la distribuzione degli insediamenti, o quella che si potrebbe definire come "l'occupazione sociale dello spazio" e la sua relativa evoluzione storica, siano un tema meno interessante o, in ogni caso, non meritevole di uno studio approfondito.

Sia ben chiara fin da subito la base teorica della presente esposizione. Questo manuale ha l'ambizione di essere d'aiuto a coloro che, in modi diversi, siano interessati alla comprensione e allo studio della strutturazione spaziale di processi umani o sociali; di rappresentare una guida a chi desideri avvicinarsi all'analisi spaziale nell'ambito di processi di indagine tesi alla comprensione dell'uomo e della società.

Ci si basa dunque sull'ipotesi che lo studio dei fenomeni sociali possa ricevere un contributo consistente dalla comprensione degli aspetti spaziali ad esso connessi. Nel fare ciò, si cercherà di allacciare diversi approcci scientifici caratterizzati dunque da un profilo prevalentemente formale. Perché non si tratta dell'applicazione cieca di metodi matematico-statistici; si tratta di esporre, soprattutto, le chiavi di collegamento tra problemi di ricerca e possibili contributi da settori metodologici diversi. Ciò fornendo al lettore gli spunti introduttivi, cercando di rendere quanto più semplice possibile il processo di lettura e assimilazione delle nozioni qui illustrate.

Lo sfruttamento reale dei risultati analitici va però di pari passo con la loro comprensione. Ed è per questo motivo che spesso è richiesto un certo rigore nell'esposizione dei metodi e della loro evoluzione storica. Ad esempio, non credo che il lettore dovrà mai confrontarsi nella realtà pratica con l'applicazione diretta delle formule matematiche qui descritte: fortunatamente di questo si occupano oggi le applicazioni informatiche a disposizione dei ricercatori; è lecito però ignorare la logica dei metodi e affidarci ciecamente ai

[7] Pattison 1964, pp. 211-212.
[8] Hodder 1982.

risultati ottenuti dai GIS? Non conoscere non solo la loro logica, ma anche la storia di una qualsiasi metodologia significa non essere in grado di applicarla ai propri problemi di ricerca. Farlo affidandosi esclusivamente al supporto del *software* significa non essere in grado di interpretarli e non essere in grado di interpretare i propri dati significa a sua volta non essere in grado di esporli o, quantomeno, di esporli correttamente.

Conoscere le metodologie quantitative rappresenta però una necessità che va al di là della soluzione di eventuali problemi scientifici. Anche la realizzazione di una carta geografica richiede la conoscenza di certe nozioni fondamentali a livello formale. D'altronde è sempre stato così. In quest'ottica, la geografia quantitativa rappresenta un supporto fondamentale per la produzione cartografica, e non solo quella scientifica. Le carte hanno un intrinseco valore formale, ed è chiaro che colui che riesce a comunicare con linguaggi formali su temi spaziali sia poi facilitato nel costruire le proprie carte.

I "problemi" di ricerca

Forse l'unico vero insegnamento che si può dare è che l'unica cosa che conta veramente nella ricerca sono i problemi; nella scienza esistono i propri problemi di ricerca e nient'altro, ovvero le domande alle quali stiamo cercando di dare risposta. Oltre tutto, si tratta di un'indicazione vecchia e che a molti apparirà ovvia e ridondante[9]. Si desidera comunque farvi accenno in queste righe iniziali solo perché essa dovrà rappresentare per il lettore un costante riferimento per tutto il testo.

La domanda più importante per il ricercatore non è *come* questo o quello deve essere fatto. Quello che conta è capire qual è l'obiettivo della ricerca, *perché* la ricerca stessa viene condotta; oppure, più semplicemente, *cosa* si cerca di spiegare. Se non esiste un problema al quale si intende dare risposta nell'agire quotidiano non è che in fondo l'operare abbia molto significato. Questa consapevolezza è fondamentale.

L'analisi spaziale non ha molto significato all'interno di un processo di ricerca privo di problemi scientifici da risolvere. Il lettore non dovrà utilizzare questo testo dunque come una guida per risolvere i propri problemi di ricerca. Faccia attenzione al fatto che non troverà all'interno di queste pagine nè domande cui rispondere – o sulle quali impegnarsi – nè tanto meno idee per possibili tematiche o argomenti d'indagine. Le domande di ricerca scaturiscono esclusivamente da una visione soggettiva e personale legata all'esperienza dei singoli, nella quale nè questo testo nè altri potranno mai intervenire se non in

[9] Per quanto concerne la comprensione dei problemi di ricerca si faccia riferimento a Popper e Lakatos.

modo passivo e indiretto. Perciò appare inutile affannarsi nell'identificazione di tematiche specifiche nè qui nè altrove. Quello che conta è l'originalità delle visioni singolari e soggettive.

Ma nel proporre una visione soggettiva di un problema scientifico e di una sua possibile soluzione occorre utilizzare un linguaggio comune. L'utilizzo di un codice che permetta agli altri membri della stessa comunità di valutare le proprie osservazioni.

Sempre a proposito di questo, è mia personale opinione che i problemi relativi alla distribuzione spaziale siano molto simili; sia nel tempo che nelle diverse aree geografiche. Questo volume mi auspico possa essere utile a tutti coloro che pensano che, in qualche misura, le questioni scientifiche che si pongono, siano legate alla comprensione delle forme e distribuzione degli insediamenti umani, cioè a tutti coloro che pensano che la comprensione delle caratteristiche distributive delle forme di stanziamento del contesto sociale studiato rappresenti un momento fondamentale nel proprio processo conoscitivo. Alla base c'è da rispondere ad una domanda generale e comune (alla quale, peraltro mi sembra non sia stata data ancora una risposta definitiva): come si distribuiscono l'uomo e i gruppi sociali nello spazio.

L'interdisciplinarietà

Per il progresso della conoscenza, l'interdisciplinarietà scientifica dovrebbe essere considerata come un valore fondamentale. Purtroppo l'interdisciplinarità comporta rischi indiscussi per chi la persegue. A causa di diversi motivi, costituisce una strada difficile e piena di ostacoli. È un dato inconfutabile che questo approccio comporti uno sforzo addizionale. In parole semplici, è necessario avere più competenze, anche molto diverse fra loro, cercando di capire l'essenza di ogni settore coinvolto per poter creare i nessi in modo coerente fra le diverse parti. Nello stesso tempo, d'altro canto, tutti coloro che seguono la strada antitetica a quella interdisciplinare si impegneranno naturalmente nella difesa strenua della propria ortodossia; in modo quasi sempre effettivo grazie ad una maggiore competenza nel proprio ambito di ricerca. Purtroppo, dunque, chi segue la strada interdisciplinare si troverà sempre di fronte un gruppo ampio di studiosi appartenenti a più discipline pronti a esprimere critiche tutto sommato giustificate. La verità è che l'interdisciplinarietà viene promossa un po' da tutti, ma pochi si impegnano veramente nel praticarla. Se l'interdisciplinarità è un bene comune, purtroppo, non molti sanno cosa sia veramente. Oltre a questo occorre sottolineare come i tentativi di creare punti di incontro e spazi comuni tra le discipline siano sostanzialmente fallimentari[10].

[10] Hall 1992, pp. 164-165.

In generale si può affermare che l'interdisciplinarietà venga solitamente confusa con la "multidisciplinarietà", ovvero l'accostamento di più discipline. Se si vuole poi definire, con una semplificazione, un elementare mezzo per mettere insieme, ma non collegare, più visioni del mondo: un contenitore in cui raccogliere più discipline senza però mai integrarle e coglierne i legami logici. Non si tratta di avere la disponibilità e la gentilezza ad ascoltare l'altro. Non si tratta nemmeno di avere la volontà di capire l'altro. Si tratta invece di avere la capacità di interagire con l'altro nel processo di formulazione di nuove idee e concetti. Purtroppo questo accade raramente.

Il tema di questo testo è intrinsecamente interdisciplinare. Esso prevede in primo luogo l'interazione di diverse branche e sotto-branche della geografia. Analisi spaziale, geografia quantitativa, geostatistica, analisi locazionale, cartografia numerica ecc. Nello stesso modo però verranno esposte idee legate a storia, filosofia della storia, matematica, statistica e geometria.

Questo non è un testo sull'interdisciplinarietà, ma attraverso appunto un discorso interdisciplinare si cerca di costruire un quadro più ampio sull'analisi dei dati spaziali, per una, per quanto possibile, corretta comprensione della geografia umana. Con tutti i vantaggi, ma soprattutto con tutti i difetti che questo può comportare. Ma è solo attraverso un percorso interdisciplinare che le problematiche di ricerca di cui si diceva sopra possono essere veramente rese più nitide. Se è vero che l'approccio interdisciplinare permette di estendere notevolmente la portata, ma anche la comprensione, delle problematiche di ricerca, d'altra parte è anche vero che in modo direttamente proporzionale aumenta la possibilità di sbagliare.

I numeri e i linguaggi formali

Nell'ambito dell'epistemologia, lo scambio di idee è considerato un fattore fondamentale. Comunicare è l'unico modo per far crescere la conoscenza. La scienza non sarebbe nulla senza la comunicazione. A questo servono appunto giornali, periodici, convegni e pubblicazioni scientifiche. Se non si disponesse della possibilità di riferirsi a questo o a quello, poco si potrebbe fare perché l'interlocutore acquisisca piena coscienza del significato che possiede e si vuole dare a quella cosa.

D'altronde il peso e l'importanza di un approccio formale o statistico nella storia è ben noto da tempo: «Un altro punto che ha bisogno di qualche chiarimento, è l'ampio ricorso all'indagine statistica. Molti troveranno eccessivo lo spazio dedicato [...] a nude colonne di numeri. Sia ben chiaro che io non attribuisco alcun valore magico alle cifre in se stesse. In molti casi esse rappresentano una semplice fonte, con problemi di esegesi suoi propri, al pari di

una cronaca o di un cartolaio.» Queste parole di Elio Conti, della prefazione di *La formazione della struttura agraria moderna del Contado Fiorentino*[11], si adattano perfettamente allo spirito e agli obbiettivi di questo volume.

Il problema però non è tanto quello di disporre di canali di comunicazione, quanto quello di utilizzare il migliore linguaggio per descrivere nel migliore dei modi l'oggetto o il processo che si sta studiando[12]. Il linguaggio da utilizzare dipende sempre dal concetto che si vuole trasmettere. Nel trattare i processi di distribuzione spaziale dei fenomeni sociali, occorre ricorrere a forme di comunicazione che consentano lo scambio preciso delle idee.

La questione che qui si desidera mettere in evidenza è semplicemente quella relativa a quale sia il linguaggio più adatto a disquisire sui fenomeni di tipo spaziale dell'uomo. Il problema è evidente in quanto, per diversi settori delle scienze umane, alcune scienze sociali e la storia, non vi è un protocollo comune con il quale formulare concetti e giungere allo scambio soddisfacente dei medesimi. In questo senso, l'assenza di linguaggi comuni rappresenta, a prescindere dai buoni propositi[13], un limite concreto all'interdiciplinarietà alla quale si alludeva sopra. È fondamentale che specialisti appartenenti a più discipline sull'uomo, nello scambio reciproco di idee, siano in grado di fornire e convogliare le proprie idee in modo preciso.

A prescindere da eventuali tentativi di formulare nuovi schemi di comunicazione condivisa, non esiste per ora alternativa ai linguaggi formali. Spesso si ha la sensazione che settori della conoscenza non (o non ancora) inclini – per dirla in modo semplice – alla statistica o al linguaggio matematico, diano un peso criptico a queste forme di comunicazione. D'altronde è naturale: si considera criptica qualsiasi forma di codificazione non nota.

Uno dei messaggi che questo testo desidera mettere in evidenza è che, in fondo, le parti del linguaggio matematico necessarie a formalizzare le idee relative alle forme d'occupazione sociale dello spazio che necessitano di essere acquisite per lo scambio di idee all'interno della comunità scientifica sono, tutto sommato, in numero limitato. Gran parte dei processi a livello delle scienze umane rientrano esclusivamente tra quelli di natura descrittiva. Non vi è migliore modo di descrivere che farlo in modo condiviso.

[11] Conti 1965, p. ix.

[12] Wicken 1987, p. 181: «Yet expressions such as "information content" are at best asymptotes for meaning. Simple, codified messages may be conveyed unambiguously; but for the most part, language is a mechanism to communicate meanings across very noisy channels. Sometimes the messages are revised and improved in response to receiver-feedback. Einstein, with characteristic modesty, claimed he didn't really understand relativity until Eddington explained it to him.»

[13] Per buoni propositi intendo l'idea introdotta da Popper con *Il mito della cornice* in Popper 1995b, pp. 57-94.

In prima istanza è evidente che l'utilizzo di un tale linguaggio rispecchia un'esigenza tipica di un approccio di tipo quantitativo. La scelta però di dare molto spazio ai numeri non vuole rispondere esclusivamente alla necessità di esporre formalmente quello che di fatto lo è già[14]. In ultima analisi, il linguaggio matematico e quello naturale sono (come peraltro tanti altri) sistemi di scambio di idee. Nello stesso tempo, è evidente che l'applicazione di linguaggi formali presenta esplicite difficoltà in un settore poco incline o scarsamente avvezzo al suo utilizzo. Si tratta di ostacoli concreti.

La scelta di impiegare numeri e formule risponde soprattutto alla volontà di favorire, in modo chiaro e diretto, scambio e diffusione di idee e di informazioni raccolte, relative a processi o fenomeni umani e sociali nello spazio. Risultato che evidentemente non sarebbe possibile conseguire con i metodi tradizionali di descrizione relativi agli studi sulla storia dell'insediamento. In altre parole, le difficoltà riscontrabili nell'utilizzo del linguaggio matematico sono compensate e ripagate dall'acquisizione della possibilità di delineare in modo puntuale e concreto concetti, idee e misurazioni sul paesaggio umano. A che servirebbero obiettività e rigore nelle indagini se parallelamente metodi di comunicazione e scambio delle informazioni relative alle realtà sociali non possedessero chiarezza e precisione?

Se un ruolo della geografia è, fra gli altri, quello di descrivere lo spazio umano, è lecito cercare metodi migliori per farlo, intendendo con tale aggettivo più precisi, cioè più efficaci nel raggiungere lo scopo prefissato, senza con questo togliere nulla ad altri modi di dialogare, per così dire, più tradizionali.

Oltre a ciò va considerato un altro aspetto di grande importanza: i linguaggi formali costituiscono l'unico modo utile alla formulazione corretta delle misurazioni. Il calcolo della misura e la sua comunicazione costituiscono un elemento unico nei processi di costruzione di nuova conoscenza. Come verrà esposto nelle prossime pagine, gran parte delle osservazioni e delle nozioni della geografia umana possono essere tradotte in misure.

Lo spiega bene Conti quando, esaminando le motivazioni di un suo approccio statistico, afferma: «quasi tutti i rapporti economici e sociali possono essere espressi in rapporti numerici. In questo senso, la statistica rappresenta per lo storico non solo una scienza ausiliaria, ma anche una forma di linguaggio, ostico ai più ma non privo di una sobria eleganza. Una serie di cifre, convenientemente analizzate nella loro provenienza e nel loro significato, può

[14] Ad esempio, si pensi alle tecniche cartografiche moderne. È ben noto che le carte geografiche costituiscono di per sé costrutti simbolici fondati sull'utilizzo di funzioni matematiche utili alla proiezioni.

spesso riassumere lunghe ricerche e costituire la base più oggettiva possibile per caratterizzazione e confronti»[15].

Tra i principali problemi di chi si avvicina per la prima volta, in ambito universitario o professionale, ai metodi quantitativi non è solo la natura ermetica del linguaggio stesso. Questione di fondo è anche dove andare a raccogliere ed imparare le nozioni di base. La matematica offre infatti diversi modi per ottenere la stessa cosa. Ad esempio, le operazioni $\sqrt{2}$ e $2^{\frac{1}{2}}$ sono equivalenti; così, lo studente dei corsi umanistici, oltre alle ovvie difficoltà iniziali, non trova nel vasto corpo dei manuali a disposizione una coincidenza di metodologie, ma formule scritte con notazioni diverse e lacune espositive. Molti testi di geografia quantitativa danno per scontate solide basi in statistica descrittiva e inferenziale. Questo costituisce un limite, spesso insormontabile, per chi si avvicina per la prima volta a questa materia.

Il calcolatore e il *software*

Nel porsi delle domande, il ricercatore finisce per aprire nuovi orizzonti, molti dei quali spesso inesplorati. Sappiamo bene che l'esplorazione di nuovi confini rappresenta un processo che richiede l'introduzione di tecniche e strumenti nuovi. Verità è che, da una prospettiva geografica, questa attività non riguarda solo il sopralluogo di terre sconosciute, ma anche l'esplorazione di nuovi segmenti e porzioni della conoscenza: l'esplorazione di nuove idee e concetti. Anche in questo caso si rende indispensabile la disponibilità di nuovi strumenti. È sufficiente pensare alla storia della scienza e all'invenzione di strumenti quali il telescopio, il microscopio o la macchina fotografica, per citarne solo alcuni.

Nello stesso modo il ricercatore spesso si trova di fronte a quesiti e questioni per i quali non esistono o non sono state fino a quel momento realizzati strumenti di analisi. Così, il microscopio finisce per aprire nuovi orizzonti; prima le cose non visibili a occhio nudo potevano solo essere immaginate[16].

[15] Conti 1965, p. x.

[16] Bellone 2000: «La questione centrale però, non riguarda l'arricchimento sensoriale determinato da certi strumenti. La questione centrale ci ricorda che, in assenza di strumenti particolari, gli organi di senso e i pensieri, da soli, non avrebbero mai scovato i dettagli dei capillari o l'anomalia di Saturno [...] Non solo osservare ma misurare. E, anche a questo proposito, nessun apparato teorico allora esistente poteva di per sé, mettere Galilei nelle condizioni di costruire pensieri e di dedurne proprio quelle valutazioni quantitative sui rari e sui periodi delle "stelle medicee".»

La storia della conoscenza realizza un grande passo in avanti con la costruzione del calcolatore elettronico[17]. Così come il microscopio ed altri strumenti finiscono per estendere i confini della conoscenza, il calcolatore elettronico diviene lo strumento capace di compiere l'esplorazione di porzioni della conoscenza per le quali solo delle ipotesi potevano essere fatte. Il calcolatore determina la trasformazione della ricerca scientifica e con essa anche lo studio delle distribuzioni spaziali. Gran parte degli schemi metodologici illustrati nei prossimi capitoli sono stati realizzati prima dell'introduzione dei calcolatori elettronici, e alcuni molto prima; si pensi ad esempio alle trame di Dirichlet. È solo grazie all'introduzione del computer però che molte delle idee, ed in particolare la loro applicazione a problemi specifici, possono essere effettivamente concretizzate. Prima, benché formalizzate, restavano nel campo della teoria: inapplicabile – se non a piccoli campioni – senza l'aiuto di strumenti di calcolo robusti. Nelle decadi del Sessanta e del Settanta, molta della produzione scientifica riguardava alcuni metodi trattati nelle prossime pagine. La differenza tra allora ed oggi è che la maggior parte dei ricercatori non aveva, salvo rari casi, accesso ad un calcolatore elettronico.

Il calcolatore elettronico, ed insieme ad esso il *software*, devono essere considerati come gli strumenti di calcolo essenziali, senza i quali molti dei percorsi metodologici qui proposti sarebbero molto difficili se non impossibili. L'analisi spaziale, come sotto-branca della geografia, si diffondeva (prematuramente) quando la disponibilità di calcolatori elettronici era tutto sommato molto limitata per non dire inesistente. La decade degli Ottanta portò con sè la rivoluzione dei *software* GIS, ma rimanevano ancora inaccessibili a molti dei ricercatori, sia per la loro complessità che per i loro costi. La successiva comparsa, durante gli anni Novanta, dei cosiddetti *desktop* GIS avviene al di fuori di una logica scientifica: sono gli anni d'oro del *software* commerciale.

Vi è un punto nel quale il cerchio si deve chiudere. Quello in cui i ricercatori nella enunciazione dei propri problemi di ricerca, attraverso l'acquisizione coerente di un linguaggio formale riescono finalmente ad approfittare del calcolatore elettronico per risolvere delle questioni a livello empirico. Naturalmente ciò rappresenta un percorso difficile, ma l'unico al quale si possa concretamente pensare per restare all'interno di un quadro sufficientemente rigoroso. Pur ricordando che l'utilizzo del calcolatore costituisce il mezzo ma non il fine.

Nello scrivere queste righe desidero mettere il lettore in guardia riguardo a certi aspetti: la comparsa del calcolatore non costituisce però un rimedio universale, anzi. Si tratta in fondo di un aspetto tecnologico ma non scientifico. In campo informatico è ben nota l'espressione *garbage in, garbage out*: di per

[17] La sua storia è alquanto lunga e non desideriamo ripercorrerla in questa sede. Si consiglia comunque la lettura di Middleton, Wardley 1990.

sè il calcolatore non conferisce un significato diverso ai dati di partenza. Esso rappresenta solo la strada obbligata dalla quale passare per risolvere, come è stato più volte segnalato, certi tipi di quesiti caratterizzati da un livello di complessità spaziale maggiore.

1
Ricerca storica e geografia quantitativa

1.1 Introduzione all'analisi spaziale

Le scienze dell'uomo si occupano in generale dello studio di processi che hanno risvolti a livello spaziale. Questo perché l'uomo è anche un agente spaziale[1]. Di conseguenza, la società e la cultura si concretizzano prevalentemente in processi con un marcato aspetto spazio-temporale. «Whatever occurs, occurs in space and time. Therefore our perception of the world is inherently spatial and temporal»[2]. La storia, e con essa tutte le discipline umane e sociali, non sono certo un'eccezione a questo. La narrazione storica ha da sempre a che fare con eventi e processi che avvengono all'interno di una cornice delimitata dal tempo e dallo spazio.

Anche se tale ragionamento appare subito impregnato di una logica e di una razionalità ben evidenti ed esplicite, è innegabile che le discipline dell'uomo, o della società, faticano e trovano molte difficoltà nell'utilizzo o nella costruzione di strumenti critici relativi alla dimensione spaziale; ovverosia nell'includere il contesto spaziale nei processi d'indagine tesi alla comprensione dell'uomo. Se per le scienze sociali lo spazio è una costante, quasi sempre è messa in secondo piano, spesso trascurata nelle forme espositive, nell'identificazione del suo ruolo nei processi e dinamiche sociali così come nella sua evoluzione cronologica. Conseguenza di tutto questo sono le difficoltà che le scienze umane e le discipline storiche trovano nell'applicazione, e sfruttamento, del concetto di spazio. Tale problematica è però scomposta in due questioni ben diverse. La prima è riferita alle cause che hanno portato ad escludere dalla scienze dell'uomo – ad eccezione di rari casi – lo studio formale degli eventi e delle cose sotto una prospettiva spaziale. La seconda questione è se questo disinteresse

[1] Sul tema si potrebbe inserire una vasta serie di riferimenti provenienti da più discipline. In questa parte si farà riferimenti in particolare a Foucault, del quale autore, per il rapporto uomo-spazio, si consiglia la lettura di Foucault, Miskowiec 1986, pp. 22-27.

[2] Wegener 2000, p.3.

Fig. 1.1. L'importanza della "posizione". Osservando questa fotografia – castello di Murlo (SI) – appare evidente come la "posizione" giochi un ruolo fondamentale. In questo caso, ad esempio, è possibile osservare come la distribuzione delle unità abitative all'interno del villaggio definisca le loro caratteristiche funzionali a livello urbanistico. Nello stesso modo, la funzione dell'insediamento all'interno del paesaggio è determinata dalla sua posizione nell'ambito della maglia insediativa.

sia più o meno lecito. Detto in altre parole, se le discipline storiche e sociali nel loro complesso siano state legittimamente portate (da una riflessione critica o meno) a considerare del tutto superfluo l'approccio formale allo spazio nel processo di formulazione di nuovi piani della conoscenza[3]. Nell'ambito della geografia umana diviene molto difficile non interrogarsi a questo proposito dato che tale disciplina, per sua stessa natura, si colloca irrimediabilmente a ridosso del crocevia interdisciplinare delle scienze dell'uomo.

[3] Per limiti di spazio è impossibile trattare qui questo aspetto; è ben evidente però che un dibattito concreto sulla "legittimità" non ha mai avuto luogo, se non attraverso casi sporadici ed episodici tra esponenti, in alcuni casi anche importanti. Vedi Schaefer 1953; Hartstorne 1954; e Hartstorne 1955. In particolare in Schaefer si può leggere come: «The scientific view, which is here taken, claims that all the data, which the historian [...] collects, are nothing but raw material for the social scientist. The historian, in constructing his pattern is, therefore, whether he knows it or not, a social scientist. [...]»

Effettivamente potrebbe darsi che una metodica di tipo spaziale all'interno dei processi di indagine umana e sociale sia del tutto inutile[4]. Questa, però, non è l'unica spiegazione di una possibile legittimazione a non considerare tale prospettiva. Come sarà esposto – e si cercherà di dimostrare – di seguito, l'approccio aspaziale delle discipline storico sociali verso i loro oggetti di indagine, potrebbe dipendere invece da stadi ancora prematuri nell'evoluzione dei contesti di studio e di conseguenza degli apparati metodologici adottati.

Con il concetto "aspaziale" non si intende solo l'assenza (parziale, o in alcuni casi totale) di una sensibilità e attenzione per lo spazio. Non contemplarlo affatto risulterebbe infatti quasi impossibile. D'altronde le carte e le mappe di distribuzione costituiscono da sempre corredo delle opere nell'ambito della storia e delle scienze umane. Quello che si intende per atteggiamento aspaziale è il non conferimento del giusto o corretto peso e significato allo spazio in termini di costruzione di nuova conoscenza. Intimamente legato a questo aspetto, ed in parallelo con l'impostazione di quanto detto, vi è la questione della comprensione di cosa sia una carta geografica ed in particolare di quale siano il suo ruolo e la sua funzione nei processi di ricerca scientifica[5].

Nella maggior parte dei casi, nelle scienze dell'uomo, alla carta geografica viene riservato infatti un ruolo prettamente passivo. Lo spazio, e le strutture descrittive, vengono così percepite come uno schema chiuso. Si tratta solitamente di una ricostruzione di tipo materiale, in cui la realtà appare come un quadro o fotografia inerte e dove le dinamiche e le interazioni spaziali a livello singolo o di sistema hanno un ruolo secondario se non nullo.

In questo senso lo spazio, e con questo le caratteristiche distributive degli insediamenti, acquistano un ruolo marginale. Non sono poche le discipline, fra le quali la stessa geografia, che usano di frequente la metafora del palcoscenico. Lo spazio appare come una sorta di palcoscenico generico all'interno del quale i processi vengono portati a termine. Ovvero, secondo l'approccio aspaziale, gli eventi accadono in un certo luogo a prescindere dalla posizione o distribuzione. In base a questa logica, se l'evento sociale viene considerato come un fenomeno appartenente ad una dimensione astratta, è lecito considerare il luogo dove esso

[4] Per una panoramica sulle questioni legate allo studio dei gruppi sociali sotto una prospettiva spaziale si consiglia la lettura di Buttimer 1969.

[5] In termini epistemologici il tema è molto articolato dato che porta direttamente alla questione dei "modelli". Per una visione precisa della questione si rimanda a Turnbull 1996, p. 6: «[...] cartography and science have been portrayed as having a closely interwoven history. The most common metaphor for scientific knowledge or theories was the map, [...] The development of "scientific maps" was taken to be identical with a progressive, cumulative, objective and accurate representation of geographical reality and, hence, was also assumed to be synonymous with the growth of science itself.». Consultare anche Dent 1972.

avviene come un fattore secondario. Perciò non meritevole di attenzione o non sufficientemente interessante per essere analizzato[6].

Sotto una prospettiva storica, il problema acquista ancora maggiore importanza. Di fatto, al di là delle diverse considerazioni, partiremo dal presupposto che uno degli elementi più importanti nella definizione di una struttura spaziale sarà la sua storia, qui intesa come evoluzione cronologica del fenomeno. Se si considera la Storia come la disciplina dell'accaduto[7], è anche vero che questo accaduto non contraddice la teoria dei sistemi. In particolare nella storia dell'insediamento, il peso del passato (qui inteso come *feedback*) può essere considerato fondamentale per la sua comprensione. Naturalmente lo studio dell'evoluzione storica dei fenomeni di stanziamento comporta un livello di complessità ulteriore; fattore che implica inevitabilmente un livello di rigore ancora superiore nella fase d'analisi.

Non si può costruire la storia di un gruppo sociale (a dire il vero, forse, nemmeno quella di un singolo) al di fuori della dimensione spaziale senza costruire una falsificazione della realtà. La sequenza cronologica degli eventi è fondamentale; ma il dove lo è altrettanto.

1.2 L'analisi spaziale e la ricerca scientifica

Per molte discipline, in particolare quelle scientifiche, è sorta da molto tempo la necessità di comprendere in modo più accurato e approfondito le caratteristiche spaziali relative ai fenomeni studiati. Sicuramente questa necessità parte da un crescente desiderio della scienza ad iniziare la comprensione e lo studio dei fenomeni nella loro complessità.

In questo senso, un'apertura alle discipline geografiche (geografia si intende qui come scienza dello spazio, o quanto meno scienza della distribuzione dei fenomeni nello spazio) può senza alcun dubbio fornire a qualunque approccio una visione più ampia sul tema di studio. Non vi è alcun dubbio che la diffusione di un approccio geografico-quantitativo nella ricerca storica dovrebbe avvenire in un'ottica interdisciplinare e non multidisciplinare.

L'approccio interdisciplinare non è una parola d'ordine né tanto meno una moda. Esso costituisce una necessità universale per tutte quelle discipline che oggi intendono procedere verso livelli di conoscenza più approfonditi; un requisito necessario per avvicinarsi ulteriormente ad una comprensione più or-

[6] In questo senso la costruzione e l'utilizzo della carta geografica costituisce un elemento fondamentale del ragionamento sull'uomo. Ora, prima di procedere alla costruzione di carte più efficaci o all'analisi delle medesime, è fondamentale cogliere e assimilare la funzione dello spazio nelle dinamiche sociali.

[7] Vedi Bloch 1969: « "Scienza degli uomini", abbiamo detto. [...] Bisogna aggiungere: "degli uomini nel tempo".»

ganica. Questo era e rimane uno dei progetti incompiuti di una storia sensibile e attenta a certe questioni: «Una crisi generale travaglia le scienze dell'uomo, sopraffatte come sono dai loro stessi progressi, non fosse altro che a causa dell'accumularsi di nuove conoscenze e dalla necessità di un lavoro collettivo, la cui organizzazione razionale rimane da realizzare»; cosí Fernand Braudel apriva il suo saggio *Storia e Scienze Sociali, la «lunga durata»*[8]. Tale progetto resta, purtroppo, a tutt'oggi incompiuto[9].

In questo senso, l'incorporazione di una geografia quantitativa al potenziale metodologico delle scienze dell'uomo non costituisce affatto un'operazione di facciata, bensì una profonda trasformazione dell'impalcatura teorica dei quadri d'azione del procedere scientifico di questa o quella disciplina. Significa per prima cosa iniziare a cogliere i fenomeni e la loro evoluzione storica all'interno di una inquadratura di tipo spaziale. Ma, a prescindere da queste considerazioni, la questione dell'applicazione dell'analisi spaziale all'interno della ricerca scientifica va al di là del dibattito interdisciplinare. La sua necessità è e resta improrogabile.

Ciò che qui si desidera segnalare è che le "caratteristiche distributive" rappresentano uno dei principali segni distintivi di qualsiasi fenomeno o processo umano e, ancora di più, sociale. Tale segno distintivo, qui inteso come attributo, sarà chiamato di seguito, anche "struttura spaziale". Si pensi ad esempio alle stelle. È del tutto lecito giungere alla conclusione che la struttura spaziale delle stelle (cioè, come queste sono effettivamente disposte fisicamente nello spazio) rappresenti una delle caratteristiche dell'universo, come d'altronde dimostrano le linee di ricerca attuali dell'astrofisica. Perché, se le stelle fossero disposte in un altro modo, sicuramente le caratteristiche dell'universo sarebbero diverse. Lo stesso discorso potrebbe essere esteso a qualsiasi altro campo scientifico che abbia a che fare con piani più o meno complessi della realtà.

1.3 La distribuzione spaziale

Un bosco, ad esempio, potrebbe essere composto da un numero n di piante, e da questo risultare più o meno densamente popolato. Ma la densità (qui intesa come un indicatore di sintesi della composizione) non è tutto. Fatto indiscutibile è che le particolarità distributive delle piante rappresentano una caratteristica (un attributo) di questo bosco. Per particolarità distributive s'intende il modo in cui gli elementi che compongono un certo sistema o fenomeno (così come gli alberi compongono il bosco, o le stelle l'universo) sono effettivamente distribuiti in termini spaziali al suo interno. Il punto che qui

[8] Braudel 1982, p. 153.
[9] Burke 1992.

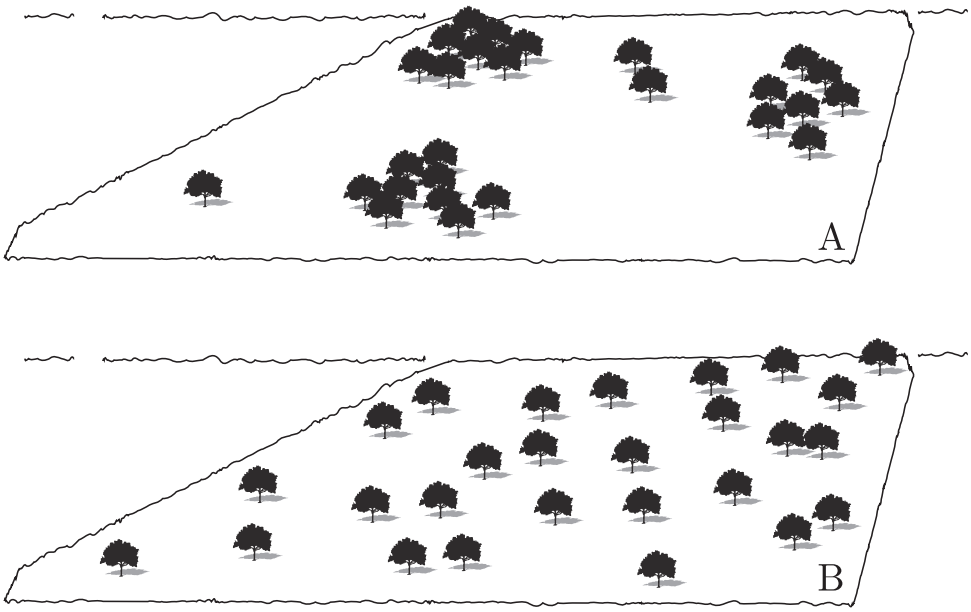


Fig. 1.2. Composizione e distribuzione. Una descrizione fondata esclusivamente sulla composizione è parziale e incompleta. La distribuzione dei componenti acquista un ruolo anche nelle caratteristiche funzionali e negli attributi dei fenomeni.

si vuole mettere in luce è il seguente: la struttura spaziale rappresenta non solo un generico elemento distintivo di un certo fenomeno, ma costituisce a sua volta una condizione intrinseca in grado di condizionare le caratteristiche funzionali dell'oggetto stesso. Per funzionali si intende proprio il modo in cui questo bosco "funziona": come risponde a condizioni esterne e come modifica il contesto di altri fenomeni.

Così la distribuzione delle piante all'interno del bosco non solo lo distingue fisicamente da altri boschi, ma fornisce indicazioni sulle caratteristiche relative al suo funzionamento. Perché la struttura distributiva delle piante è causa ed effetto di altri processi che avvengono al suo interno. In tutti questi casi il tema centrale potrebbe non apparire affatto direttamente legato alla struttura spaziale. Nonostante ciò è facile dedurre come sia indispensabile considerare anche gli attributi distributivi (la loro struttura spaziale) delle piante nel bosco per giungere alla comprensione del fenomeno studiato e dunque alla produzione di nuova conoscenza; perché, come è stato affermato sopra, il processo cognitivo prende piede a partire da schemi deduttivi causa-effetto; ed è perciò

lecito considerare che anche la comprensione dello spazio giochi un ruolo fondamentale all'interno di questo processo cognitivo di tipo «collettivo» come quello auspicato da Braudel.

All'interno di una logica astratta e aspaziale, come può essere quella di alcuni approcci e tradizioni delle scienze umane – dove lo spazio viene trattato in modo acritico e non formale – si può effettivamente giungere alla descrizione, e quindi allo studio, di un certo fenomeno avvalendosi semplicemente del concetto di composizione. Così, dunque, un bosco è composto da 5.000 piante, una rete di stanziamento da 300 villaggi e la popolazione di un territorio da 25.000 anime. Anche se è vero che la composizione è molto importante per cogliere l'essenza dell'oggetto di studio, è anche vero che a poco serve se non si conoscono le sue caratteristiche distributive. Da qui il concetto di struttura spaziale. Cioè di come gli elementi che compongono un fenomeno sono strutturati (o distribuiti) a livello spaziale.

Per capire il senso di queste affermazioni si osservi la figura 1.2 dove sono rappresentati due boschi. Entrambi coprono la stessa superficie, ed entrambi sono composti dalla stessa quantità di alberi: 27. Se si ragionasse esclusivamente in termini di composizione, i due casi apparirebbero molto simili se non uguali. Ma le caratteristiche distributive (o la struttura spaziale) delle piante all'interno delle due aree sono evidentemente diverse. Il bosco B presenta un indice di dispersione maggiore rispetto a quello A. Queste caratteristiche distributive non hanno un ruolo secondario; certamente non inferiore a quello della loro composizione. Eppure gli aspetti distributivi sembrano godere da parte degli studiosi di una considerazione minore. Ci si impegna a enunciare la composizione dei fenomeni, ma non la loro articolazione a livello spaziale.

Per dimostrare come la distribuzione non sia affatto secondaria, si pensi a quanto segue. La struttura spaziale delle piante condizionerà, ad esempio, il rischio di incendi, o la velocità e la possibilità di propagazione delle malattie al suo interno. Nello stesso modo le due distribuzioni di alberi avranno un effetto diverso nell'erosione del suolo. Da questo si può facilmente dedurre come gli effetti della distribuzione abbiano anche un ruolo condizionante molto importante sulla stessa vita futura del bosco.

Si potrà così facilmente concludere come, in linea teorica, la struttura spaziale del bosco A possa favorire, in virtù del maggiore grado di dispersione delle piante, la diffusione delle malattie. Lo stesso discorso vale anche in caso d'incendio. Se l'incendio investisse il bosco A è facile immaginare che questo rimarrebbe isolato in una delle tre concentrazioni, mentre si potrebbe diffondere liberamente nel caso di quello B. Le caratteristiche funzionali sono diverse se si pensa ad esempio all'erosione. In questo caso le caratteristiche distributive del bosco A favoriranno l'erosione delle ampie aree non coperte da piante.

Perciò è fondamentale ricordare sempre come, nella definizione o raffigurazione delle caratteristiche di un fenomeno da sottoporre a una fase d'indagine scientifica, l'idea relativa alla composizione conti relativamente poco se essa non è accompagnata da quella della distribuzione.

La struttura spaziale dei fenomeni è proprio questo: il modo in cui le parti che compongono un fenomeno si organizzano a livello spaziale. Definizione che acquista ancor maggiore valore se il fenomeno sotto esame è un sistema. L'unico modo per far funzionare una macchina, o un organismo, è una precisa distribuzione e posizione delle parti. Non è sufficiente, e forse non è possibile spiegare i fenomeni in modo scientifico senza una nozione della loro disposizione a livello spaziale.

Tra le diverse variabili che determinano la nascita, l'evoluzione e le caratteristiche dei fenomeni, quelle relative alla loro natura spaziale acquistano un peso sempre più elevato. Alcune discipline come l'ecologia, la zoologia, la geografia, l'economia e l'archeologia sono in parte giunte alla conclusione che, per arrivare a forme di conoscenza reale all'interno dei processi di indagine, non potevano essere trascurati tutti quelli aspetti o variabili relativi allo spazio. L'importanza crescente che le variabili spaziali acquistano all'interno dei quadri e processi di indagine correnti può dipendere dalla difficoltà di allargare ulteriormente i confini della conoscenza tracciati con metodi convenzionali. Messa sotto un'altra luce, la questione potrebbe essere letta nei seguenti termini. L'utilizzo di un numero limitato di variabili ha portato all'apprendimento complessivo di una parte del sistema, di una porzione limitata del quadro sotto indagine. Per un allargamento della visione complessiva del quadro o del sistema studiato, si rende necessaria l'introduzione di nuove variabili fra gli argomenti di ricerca; e non solo, visto che quello che occorre veramente è l'incorporazione (e correlazione) di queste nuove variabili all'interno delle strutture di conoscenza già costituite. Se la distribuzione e la conseguente struttura spaziale sono così importante per oggetti come stelle o essere viventi inanimati come le piante, lo è ancora di più per i gruppi sociali. Ed è per questo motivo che lo studio e la comprensione di tali *patterns* risultano così importanti[10].

1.4 L'importanza della posizione nella ricerca sull'uomo

Oltre alla struttura spaziale di un determinato fenomeno sociale è fondamentale anche la comprensione della "posizione". Di questo aspetto si è occupata prevalentemente una parte della geografia economica. Essa «si presenta come una fisica sociale e tenta di scoprire come si dispongono recipro-

[10] Rushton 1969, p. 391.

camente gli atomi e le molecole del corpo sociale e quali siano i legami che li caratterizzano. . . »[11].

È chiaro che l'analisi della posizione dei luoghi può avere grande importanza e peso in processi d'indagine relativi alle scienze naturali. Ma quanto sia importante nell'ambito della ricerca sull'uomo non è immediatamente intuibile. Nonostante la psiche umana sia costantemente immersa all'interno di una dimensione spazio-temporale, è molto difficile cogliere l'importanza della dimensione spaziale[12]. Basti comunque pensare che, senza lo spazio, il movimento o la sopravvivenza degli esseri animati non sarebbe possibile. Dunque si parte dall'idea che, in ogni caso, lo spazio è una condizione indispensabile per l'essere umano e per tutte le sue attività concrete e quotidiane. Di fatto non è necessario solo lo spazio per il movimento dell'individuo, ma anche quello per la produzione e la raccolta delle risorse necessarie al suo sostentamento. Si può concludere dunque che, dalle attività primarie necessarie alla sopravvivenza della specie a quelle più complesse delle società più avanzate, lo spazio rappresenti un elemento fondamentale.

Il comportamento e le scelte dell'essere umano, ma ancora di più i gruppi sociali, definiscono nella propria quotidianità una vasta gamma di *patterns* spaziali[13]. Alcuni dei quali caratterizzati da resti materiali, mentre altri da segni intangibili. I processi comportamentali, i resti, le tracce, la loro percezione può essere capita solo attraverso una chiave spaziale; sia essa formale o qualitativa, intuitiva o induttiva.

Per comprendere in modo più dettagliato l'importanza dello spazio nei processi sociali, è utile procedere alla presentazione di un esempio. Come premessa si sottolinea come il seguente schema risulti essere estremamente semplice e didascalico; nonostante ciò si potrà osservare come lo spazio finisca per condizionare in modo irrimediabile le scelte dell'uomo. S'immaginino due insediamenti: A e B. I due siti saranno qualificati come centri di una certa importanza. Per centri s'intende che essi sono caratterizzati da una popolazione consistente, un mercato e altre strutture di servizio[14]. Si dovrà inoltre immaginare che essi siano posti a una distanza consistente l'uno dall'altro, restando sempre e comunque all'interno di un unico ambito territoriale. In generale si

[11] Claval 1983, pp. 195-196.

[12] Il tema potrebbe apparire secondario o puro strumento retorico. In realtà l'argomento è centrale nella logica della presente esposizione. Per una panoramica generale sulla questione spazio ed esseri animati si rimanda addirittura a una figura come Poincaré 1963, pp. 15-16: «I shall exclude first of all the idea of an alleged sense of space which would locate our sensations in a ready-made space whose notion would pre-exist all experience, and which prior to all experience would have all the properties of the space of the geometer[. . .].»

[13] Flannery 1976.

[14] Vedi modello di Christaller; capitolo 3.

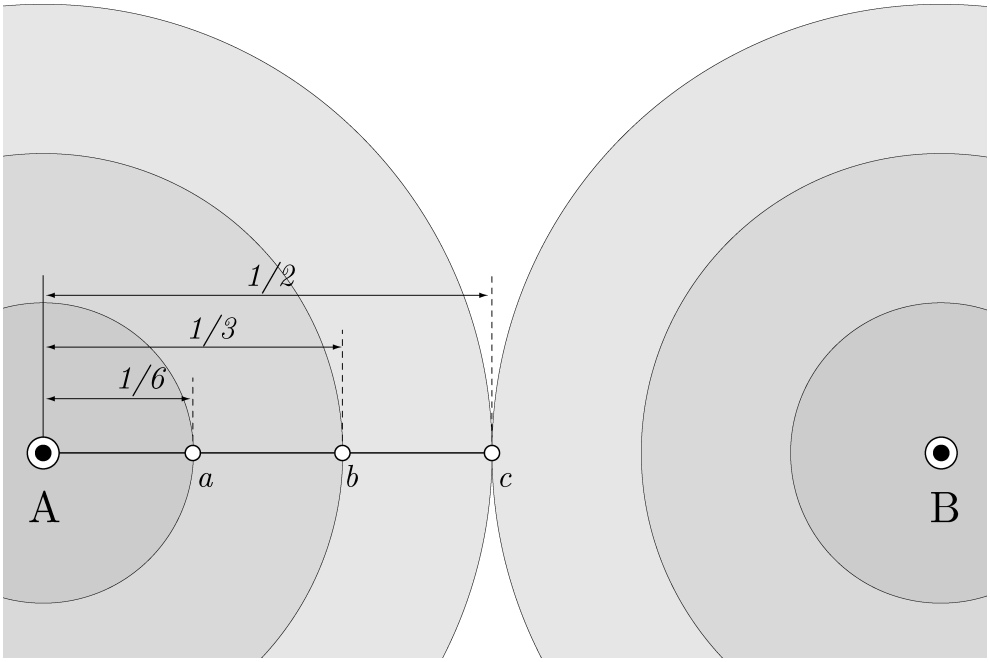


Fig. 1.3. Il ruolo dello spazio nelle dinamiche di stanziamento. Benché i tre insediamenti fra i centri A e B possano considerarsi identici, la loro posizione nello spazio determinerà e condizionerà le loro caratteristiche e le loro funzioni. Vedi Lloyd, Dicken 1972, p. 20.

potrà affermare che gli insediamenti A e B siano equivalenti; cioè essi svolgono la stessa quantità e tipologia di servizi.

Si immagini dunque che, in mezzo allo spazio tra questi due grandi centri, si collochi una serie di piccoli insediamenti. Un caso concreto potrebbe essere quello di tutti i centri sorti lungo la strada che collega appunto A a B. Ognuno di questi piccoli insediamenti è composto da pochi individui: un'unica unità familiare. E la loro funzione può essere qualificata come specializzata nella produzione agricola. In questa dimostrazione è importante sottolineare che le caratteristiche morfologiche e funzionali di questi centri siano da considerare identiche, sia per composizione sociale che per tipologia e indici della produzione[15].

[15] In realtà questo presupposto è improbabile visto che nella pratica la posizione finisce per condizionare il tipo di produzione dei piccoli centri agricoli. Vedi modello di Von Tünen.

Si prendano in considerazione tre stanziamenti posti rispetto ad A a 1/6, a 1/3 e a 1/2 della distanza tra A e B. Essenzialmente i tre centri potrebbero essere visti come equivalenti. Di fatto sono uguali (anche se, come sarà dimostrato di seguito, solo in apparenza). Tutti e tre i gruppi avranno la stessa tipologia e livello di produzione. I tre siti hanno lo stesso numero di abitanti e la stessa forza lavoro. Anche i livelli di consumo saranno, in questa dimostrazione, identici.

Si parte così dall'ipotesi che la produzione agricola servirà principalmente per il sostentamento della popolazione di ogni singolo stanziamento. Le eccedenze avranno un valore e potranno essere trasformate in capitale o scambiate solo se piazzate nel mercato dei centri principali A o B. Di conseguenza, i prodotti dovranno essere trasportati da ogni insediamento minore verso il centro A. Al costo di produzione sulle mietiture, si andrà perciò ad aggiungere il costo di trasporto. In termini elementari, il costo del trasporto può essere immaginato come il prodotto tra il costo medio per il trasporto di un certo tipo di bene e la distanza totale da percorrere. Dunque, se x è il costo di produzione dei beni agricoli e y è il costo della percorrenza di 1/6 della distanza tra i centri A e B, una volta piazzata la merce nel mercato di A, i costi di produzione e trasporto del raccolto dei tre siti sarà:

$$\begin{aligned} a &\Rightarrow x + y \\ b &\Rightarrow x + 2y \\ c &\Rightarrow x + 3y \end{aligned}$$

Le leggi della domanda e dell'offerta regoleranno un prezzo di mercato. Tale prezzo dovrà essere applicato a tutte le merci di stessa tipologia e qualità, a prescindere dalla provenienza a, b o c. Se si definisce z come prezzo unitario di mercato, si potrà calcolare come ricavato il prezzo di mercato meno i costi totali. Secondo l'esempio proposto, il profitto per i tre piccoli centri può essere calcolato in:

$$\begin{aligned} a &\Rightarrow z - (x + y) \\ b &\Rightarrow z - (x + 2y) \\ c &\Rightarrow z - (x + 3y) \end{aligned}$$

Non è difficile capire allora come la posizione nello spazio abbia condizionato le funzioni e le caratteristiche dei tre abitati malgrado le ipotetiche condizioni iniziali fossero identiche. Tutto questo a discapito degli insediamenti più lontani dal centro A. Infatti i loro abitanti dovranno optare per un

profitto inferiore in modo da potere concludere un numero di scambi in linea con quelli degli altri centri minori posti nelle vicinanze delle piazze di mercato. In altre parole, al momento della vendita o dello scambio, il prezzo complessivo dei raccolti sarà, per questioni di mercato, lo stesso; tutto questo a svantaggio di quei contadini che hanno percorso una strada più lunga. In altri termini, la capacità di ammortizzare i costi di produzione e trasporto dei centri più remoti sarà inferiore.

È anche vero però che i contadini dell'abitato a mezza strada (c), cioè quelli posti esattamente a metà tra i centri A e B, avranno, a differenza di quelli posti a 1/6 (a) e 1/3 (b) della distanza totale, la possibilità di scegliere valutando se esista una convenienza per loro nel distribuire i loro prodotti nel mercato del centro A o in quello del centro B. Non solo, proprio per il fatto di essere collocato a metà strada, questo stanziamento avrà maggiori garanzie nel caso si decidesse di avviare una diversificazione della produzione. Perché il mercato del centro A potrebbe giungere a presentare una domanda di prodotti più o meno diversa da quella del mercato B. Oppure, nel caso i traffici di persone o beni si intensificasse tra i centri A e B, lo stanziamento C potrebbe crescere e consolidarsi fino a raggiungere anch'esso il livello di centro, di servizio per i viaggiatori.

Nell'illustrare questo caso ipotetico si è cercato di formulare un esempio elementare proprio per rilevare come lo spazio sia determinante perfino nelle condizioni più semplici. Lo spazio non rappresenta solo un palcoscenico all'interno del quale viene rappresentata la fenomenologia dell'esperienza sociale. Lo spazio non è neppure un attore passivo. Esso costituisce uno degli agenti più forti dell'intera esperienza umana. Di fronte alla contrattazione o rapporto tra entità come spazio e cultura (o spazio e società), l'uomo, inteso qui come individuo, pur nella possibilità di scelte o reazioni alle condizioni esterne, deve, in ogni caso, adottare scelte coerenti con le condizioni dettate dalle caratteristiche distributive dei sistemi di stanziamento di appartenenza.

Nel tentativo di capire lo spazio e la logica delle scelte che gli individui sono costretti ad adottare, occorre affidarsi anche a modelli schematici come quello appena descritto. Tanto è importante questo settore d'indagine che, in seno alla geografia, nel tempo (grazie anche a una tradizione molto lunga) si è formata una branca specifica nota come "analisi locazionale" (vedi cap. 6).

1.5 La complessità spaziale

Ogni processo teso alla costruzione di nuova conoscenza passa necessariamente attraverso la fase di descrizione e comunicazione del fenomeno studiato.

In larga misura si tratta di descrivere l'idea che, sotto forma di modello più o meno articolato, si è costruito della realtà. Perché la conoscenza non è tanto quello che viene imparato dalla realtà, quanto il processo di consapevolezza sulle idee di essa acquisite e maturate all'interno di un processo evolutivo nel tempo.

Se è vero che lo spazio ha così tanta influenza sulla realtà quotidiana dell'uomo, è lecito chiedersi dunque i motivi dell'evoluzione di una ricerca sociale di tipo aspaziale. Si rende necessario cioè spiegare il perché le ricostruzioni storiche, i modelli sociali, demografici e perfino, in molti casi, anche economici, siano concepiti senza prendere in considerazione l'impalcatura spaziale dei fenomeni[16].

Chiaramente bisogna sottolineare come non sia possibile parlare in modo assoluto di una assenza o di un disinteresse, e in qualsiasi fase evolutiva delle discipline storico-sociali, verso il tema spaziale. Le eccezioni ci sono sempre state. Naturalmente per attenzione al tema dello spazio ci si riferisce a qualcosa di più che fornire una carta geografica a corredo della narrazione. Il problema sta proprio qui: spesso le cose raccontate contrastano con le informazioni che derivano o potrebbero essere desunte dai dati cartografici. Con tutto questo non si vuole suggerire un percorso tale da rendere lo spazio l'unico elemento delle ricerche sull'uomo. Si desidera solo rilevare come l'approccio aspaziale costituisca un fattore comune nel quadro generale della ricerca sull'uomo ed in particolare sui sistemi di stanziamento umani. Come è stato appena indicato, non si tratta di un costante e assoluto disinteresse. È vero invece che, anche quando traspare un'attenzione verso lo spazio, spesso ci sono limiti nella corretta comunicazione e scambio di idee.

Uno dei motivi della carenza di attenzione per lo spazio potrebbe essere individuato nella natura caotica o complessa delle strutture spaziali; complessità aggravata dalle possibili articolazioni delle strutture sociali. In altre parole, è facile intricare rapidamente e gravemente le cose quando si parla di spazio, e ancora di più se si parla di spazio umano. Nell'esempio dei due centri di servizio, le cose possono apparire molto semplici; di fatto il modello è stato ridotto al minimo. Nonostante questo, è stato possibile, in una dimostrazione così semplice, osservare come le cose tendessero, per loro natura, ad articolarsi in più piani attraverso modi sempre più intricati. Sappiamo inoltre che la realtà è invece molto più complessa dell'elementare esempio del paragrafo

[16] La questione della mancanza di attenzione allo spazio (o alle posizioni nello spazio) era già stata segnalata da A. Weber 1968, pp. 1-5: «[...] political economists have dismissed this problem of location with some general rules of local and international division of labor, etc., [...] But while problems of location have been treated by geographers primarily, Thünen is a notable exeption. There are, too, several attempts in this field. But they are insignificant when compared with the magnitude of the problem.»

precedente. Basta aggiungere un fiume e una catena di montagne, diversificare i tipi di produzione, i tipi di servizi e introdurre la distribuzione di risorse naturali o variare i tipi di suolo per rendere il tutto incomprensibile; nel senso che la totalità degli elementi e delle relazioni che regolano il sistema finirebbero per sfuggire al pur attento occhio del ricercatore. Ma i sistemi di stanziamento umano funzionano. Cioè, i mercati garantiscono lo scambio dei beni, le unità produttive garantiscono (tendenzialmente) la copertura del fabbisogno di una determinata popolazione, e gli stanziamenti garantiscono la sopravvivenza biologica e culturale della specie umana. Il che implicitamente ci indica come appunto le variabili spaziali possiedano un ruolo funzionale all'interno del sistema. Se questo non fosse vero, i sistemi d'occupazione sociale dello spazio, in linea di principio, non funzionerebbero affatto. Perciò è lecito dedurre che in qualche misura le reti di stanziamento, al di là della loro complessità, possiedano forme di ordine ed equilibrio.

A differenza dello spazio, il tempo appare come un aspetto della realtà di più facile assimilazione. È importante in questo senso non confondere il tempo con la sua percezione. Per tempo si intende una delle dimensioni fondamentali dell'universo; l'unità con la quale misurare la durata degli eventi. Un aspetto molto diverso è la percezione culturale che si può avere di esso. Tali forme di percezione possono a loro volta cambiare o subire delle trasformazioni nei diversi periodi storici e all'interno delle diverse culture. Dunque possiamo desumere che le differenze nella percezione culturale del tempo possano provocare alterazioni dei *patterns* sociali e l'introduzione di nuovi *trends*. In questo senso spazio e tempo appaiono come due dimensioni altrettanto relative. Per alcuni versi il tempo potrebbe essere concepito come una linea continua. Certamente la spiegazione dei fenomeni acquista un senso solo grazie alla corretta posizione degli eventi lungo la linea del tempo. È innegabile, in ogni caso, che per lo studioso non rappresenti un problema l'atto pratico di collocare gli eventi lungo questa linea. Lo stesso sistema di convenzioni per indicare i fenomeni nel tempo dimostra la semplicità della variabile tempo. 10 agosto 490 a.C.: battaglia di Maratona.

Il sistema di riferimento spaziale è molto più complesso. Per maggiore chiarezza sottolineo come il confronto venga qui realizzato tra tempo e spazio intesi come misure fisiche dell'universo e non come percezioni culturali delle due dimensioni. Per indicare e fare riferimento ad oggetti nello spazio vengono utilizzati i sistemi di coordinate basate su diversi sistemi di proiezione geografica. Bisogna utilizzare il plurale perché, in campo geografico, esistono di fatto diverse forme per definire la posizione[17]. Anche se una città può essere

[17] Per una storia delle coordinate geografiche consultare Keuning 1955, pp. 1-24; per una classificazione delle proiezioni fare riferimento a Tobler 1962, pp. 167-175.

indicata all'interno di una carta geografica con una sola coppia di valori x y (in modo molto arbitrario, visto che essa occupa molto più spazio di un punto a zero dimensioni), diviene più complesso indicare la posizione di altri oggetti geografici quali un fiume o una montagna. Servono cioè sequenze molto lunghe di coordinate per oggetti geografici che, peraltro, occupano lo spazio a livello tridimensionale.

Questo però non è l'unico problema. Il tema della posizione spaziale diviene profondamente complesso e caotico ogni qual volta esso viene integrato con fattori cronologici[18]. Ma c'è anche da considerare che il percorso stesso di un fiume potrebbe variare nel tempo e così la sua posizione di oggi potrebbe non essere la medesima di cento o di mille anni fa. Lo stesso dicasi delle città, o delle istituzioni che nei centri trovano luogo[19].

Non solo: la popolazione è caratterizzata da una duplice trasformazione della propria strutturazione spaziale. Da una parte i fenomeni migratori comportano lo spostamento di vasti gruppi, cosicché i flussi di migrazione determinano, a seconda delle fasi congiunturali, processi più o meno massicci di spostamento delle persone. Parallelamente a questo, la popolazione, come qualsiasi fenomeno caotico, cambia costantemente. I *patterns* riproduttivi e di fertilità variano nel tempo. Perciò, anche se solo in apparenza lo schema insediativo resta immutato, il popolamento può trasformarsi in modo marcato[20].

La comprensione dello spazio è perciò resa più ardua proprio dal fatto che i fenomeni e gli oggetti tendono nel tempo a cambiare posizione, contribuendo così alla definizione di nuovi schemi distributivi. Non si studiano tanto i *patterns* spaziali quanto la loro storia.

Uno studio e un approfondimento della strutturazione geografica dei fenomeni studiati sono rimasti al di fuori degli scopi immediati di parte delle scienze umane e sociali forse perché, come è stato affermato prima, tale ambito resta molto complesso e agli occhi del ricercatore la realtà appare sostanzialmente caotica. Si può giungere alla conclusione che la carenza che c'è stata in questo senso sia fra le altre cose giustificata da una impossibilità materiale di contrastare in modo critico la complessità che si presenta e compare di fronte al ricercatore[21].

Far fronte alle diverse difficoltà che imponeva ed impone uno studio della distribuzione spaziale di gran parte dei *patterns* sociali significa assumersi l'onere di un impegno di risorse ed energie considerevole. Tutto questo all'interno di un quadro metodologico ancor oggi in piena evoluzione.

[18] Prigogine 2006, pp. 5-6. Per quanto riguarda i problemi relativi alle variabili cronologiche in processi analitici di tipo spaziale si consiglia inoltre la lettura di Sack 1974.

[19] Si prenda il caso della Diocesi di Roselle che nel 1138 fu trasferita a Grosseto.

[20] Wagstaff 1979.

[21] Appleby 1954, pp. 79-80.

Fig. 1.4. Paleoalveo: Valle dell'Albegna (GR).

Naturalmente non è possibile affermare in senso assoluto una totale disaffezione verso il tema geografico-spaziale; almeno non sarebbe possibile farlo nei riguardi della storia. Lo storico tende in ogni caso a tracciare un quadro geografico della regione trattata. Non è difficile trovare contributi di questa disciplina forniti, dove si faccia riferimento a un territorio o a una regione specifici, della relativa carta geografica . Non solo: diversi storici, sicuramente ispirati da una sensibilità interdisciplinare, verso la materia di studio hanno assunto un approccio che può essere definito di tipo geografico[22].

[22] Per i rapporti tra storia e geografia consultare Guelke 1997. Inoltre, sull'importanza degli aspetti percettivi nella ricostruzione storica, fare riferimento a Carr 2001.

Di per sé, l'analisi spaziale (intesa come l'insieme delle discipline impegnate nella comprensione dello spazio umano) non può essere considerata una grande novità[23]. Il potenziale metodologico che verrà illustrato di seguito era in larga misura noto da tempo. Le note metodologiche che compongono questo testo non rappresentano, sotto nessun aspetto, alcuna novità, né tanto meno una rivoluzione disciplinare. La carenza applicativa di una disciplina dello spazio dipendeva esclusivamente dal fatto che portare a termine un'analisi singola richiedeva una quantità molto elevata di ore/uomo. La quasi totalità di questo tempo doveva essere dedicato a processi di calcolo e verifica delle singole operazioni. Tale costo era improponibile per molti gruppi di ricerca. Di fatto, in passato l'applicabilità dell'analisi spaziale veniva relegato a progetti di intervento e organizzazione statale, con ampi fondi e risorse finanziarie; ben al di fuori della maggior parte degli ambiti scientifici o di ricerca universitari. Queste difficoltà erano notevolmente aggravate per le scienze sociali e umane. La verità è che l'analisi spaziale non si era mai diffusa in termini applicativi, restando ristretta a gruppi specifici con particolari interessi teorico-metodologici; non tanto a causa di eventuali difficoltà tecniche, ma piuttosto del costo e tempo richiesti per la realizzazione di una singola analisi.

Nella seconda metà del XX secolo si registra la diffusione del calcolatore elettronico; nelle ultime due decadi invece il potenziamento dei processori e la loro diffusione a livello commerciale. Questa diffusione all'interno dei laboratori e centri di ricerca porta con sé il seme di una nuova rivoluzione scientifica ancora in atto. Per la prima volta i ricercatori si trovano di fronte alla possibilità di definire programmi capaci di compiere le sequenze di quantificazioni una volta realizzate con carta e penna[24]. Il computer diviene l'equivalente della macchina della rivoluzione industriale capace di sostituirsi all'operatore umano in fase di quantificazione e analisi dei dati. La geografia non è certo un'eccezione a tutto questo[25].

Molte delle elaborazioni tipiche dell'analisi spaziale e della geografia quantitativa in teoria non esigerebbero l'uso del calcolatore; di fatto queste potrebbero essere realizzate ugualmente con carta e penna. Ma, visti i lunghi tempi che l'elaborazione manuale richiederebbe, è solo grazie alla diffusione di microprocessori ad alte prestazioni che questo tipo di operazioni diviene

[23] Vedi Sack 1974, p. 439.
[24] Coppock, Johnson 1962, pp. 133-135.
[25] Claval 1980, pp. 212-215.

accessibile alla maggior parte dei gruppi di ricerca[26]. La mole di operazioni di certe analisi è così grande che, perfino oggi, alcuni processi tipici della geografia quantitativa richiedono tempi d'elaborazione particolarmente lunghi anche se compiute con gli elaboratori più veloci.

1.6 Lo scopo dell'analisi spaziale

In senso generale l'"analisi spaziale" va intesa come un settore scientifico vasto e trasversale, caratterizzato da diversi approcci metodologici di tipo statistico e topologico tesi alla comprensione e allo studio della distribuzione dei fenomeni nello spazio. Ad esempio, nel campo della geografia umana, l'analisi spaziale può essere definita come quella branca che si occupa dello studio formale della struttura spaziale delle reti di stanziamento umano, oppure della distribuzione ed organizzazione dei centri nello spazio. Questo avviene attraverso la formulazione di schemi e modelli quantitativi di diversa natura (numerica, statistica, geostatistica) che sintetizzano, in modo formale, appunto, la natura o l'essenza della maglia d'insediamento. Naturalmente questo è quello che riguarda il campo della geografia umana, o comunque l'interazione dell'uomo con lo spazio. Bisogna però ricordare che, in ogni caso, l'analisi spaziale si occupa sostanzialmente dello studio di qualsiasi forma di distribuzione a livello spaziale. Di conseguenza essa viene applicata non solo all'interno della geografia umana, ma anche nel quadro generale di altre discipline come l'astronomia, la zoologia o l'ecologia, per citarne solo alcune.

In realtà, l'applicabilità dell'analisi spaziale nella ricerca sull'uomo non si limita solo allo studio di mappe per punti, come può essere ad esempio quella di una carta di distribuzione d'insediamenti. L'analisi spaziale viene estesa anche allo studio e alla comprensione di strutture lineari, come percorsi viari o reti di comunicazione, e alla comprensione della strutturazione spaziale di superfici o aree come campi e parcelle. Essa ha un ruolo fondamentale in campi come l'organizzazione delle reti di distribuzione commerciale o la gestione di quelle del trasporto urbano. Sarebbe impossibile procedere all'organizzazione della linea di una metropolitana o degli autobus di una grande città senza il supporto di un approccio quantitativo come può essere quello dell'analisi

[26] Rhind 1977, p. 71: «Much of the work previously denoted as "statistical cartography" has involved calculations prior to the mapping; the production of frequency distributions and univariate statistics to aid decisions on selection of class intervals is one such example. Such analysis has been limited in the past because of the difficulties of making mechanical computations by hand or by machine and then manually mapping the results. Now, however, it is entirely possible to carry out some statistical exercise, map the results, carry out a more refined statistical exercise suggested by the results, map the new results ... etc.»

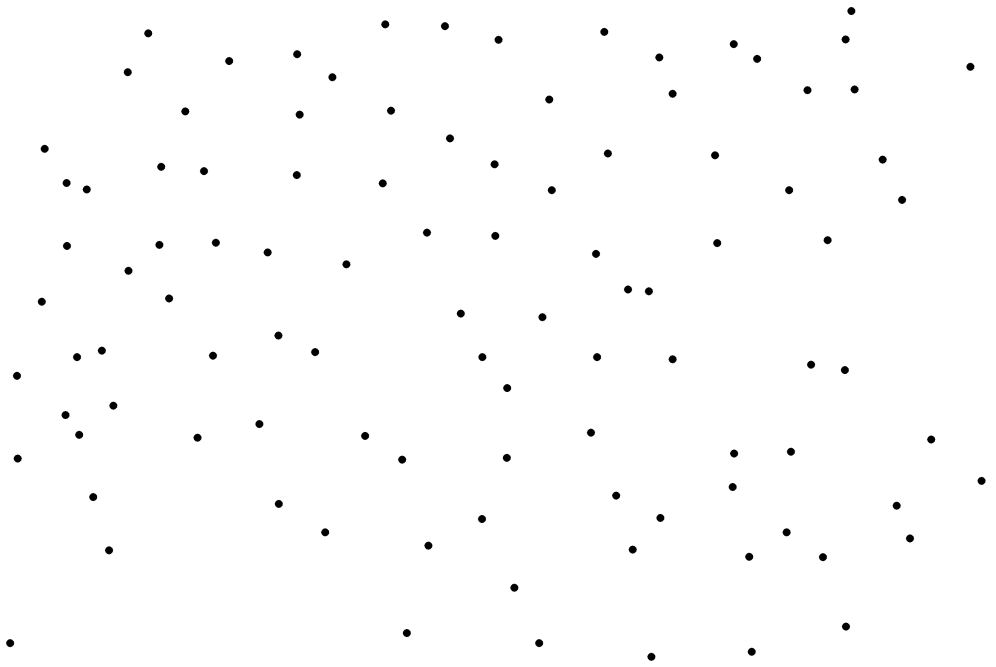


Fig. 1.5. Una maglia di distribuzione reale. La figura illustra una porzione della rete plebana della Tuscia settentrionale. A occhi nudo è praticamente impossibile cogliere qualsiasi forma di organizzazione o tendenza al suo interno.

spaziale. Lo stesso vale per l'organizzazione e la valutazione degli indici di produzione dei territori e delle regioni agricole specializzate.

Uno dei principali obiettivi dell'analisi spaziale è quello di procedere verso l'identificazione sia delle caratteristiche sommerse, sia dei *patterns* spaziali dei sistemi di distribuzione, non riscontrabili o identificabili con la semplice ispezione diretta (cioè a occhio nudo) della realtà o della loro rappresentazione attraverso le carte geografiche[27]. «Men have been drawing maps and so studying spatial patterns for millenia, yet the need to reduce such information to numbers is rather recent. The human eye form a marvelous mechanism with which to analyze and recognize patterns, yet they are subjective... »[28]. Quando si procede ad una sua prima lettura, a occhio nudo ogni carta di distribuzione appare confusa e disorganica. Ben poco si otterrebbe se fosse chiesto ad un osservatore di annotare tutte le considerazioni e le riflessioni ricavabili dall'ispezione diretta della carta di distribuzione esposta nella figura 1.5.

[27] Silk 1979, pp. 3-5.
[28] Ripley 1981, p. 1.

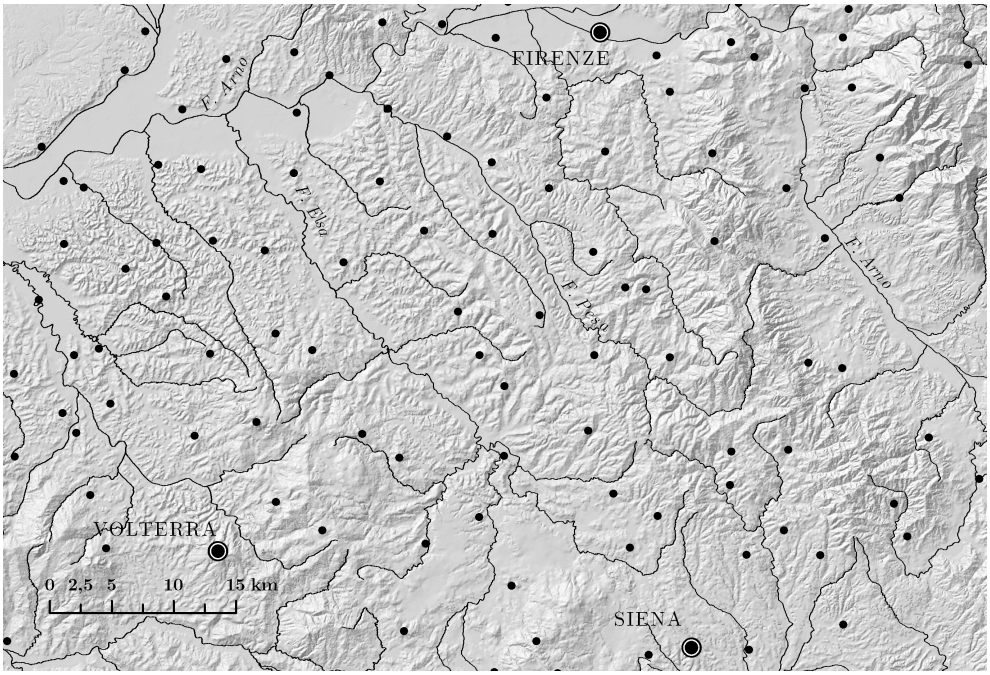


Fig. 1.6. Neppure accostando alla rete di distribuzione delle pievi il suo contesto topografico è possibile cogliere l'essenza o il potenziale d'informazione implicito in essa. In realtà l'incremento d'informazione nel piano cartografico rappresentato rende la lettura delle tendenze ancora più difficile.

Di fatto, le osservazioni si limiterebbero a fornire qualche vaga indicazione sulla densità relativa nelle distinte aree della carta: *più denso da questa parte e meno denso da quest'altra*. Purtroppo, all'interno di un processo di ricerca formale, queste osservazioni non avrebbero naturalmente alcun valore. Esse potrebbero essere archiviate semplicemente come delle semplici valutazioni di tipo soggettivo[29]. La forza del condizionamento soggettivo nella formulazione di osservazioni dirette da parte di un osservatore appare ancora più evidente se si prende in considerazione l'eventualità, molto frequente, che la rete di distribuzione analizzata possa presentare delle caratteristiche collegabili ad effetti ottici di varia natura.

Questo è il punto centrale di una parte delle questioni che il presente lavoro intende affrontare. Lo storico, l'archeologo, l'antropologo e altri ancora hanno

[29] Sulla natura soggettiva dell'osservazione e interpretazione della realtà geografica si suggerisce la consultazione di Ley 1977 pp. 503-507.

compiuto da sempre valutazioni e letture cartografiche di questo tipo[30]. Non bisogna mai dimenticare che all'interno del quadro generale delle discipline umane e delle scienze sociali si è proceduto sempre ed in ogni caso con l'ispezione diretta delle carte, senza prendere in considerazione l'immenso peso giocato dalla soggettività della valutazione dell'osservatore. Questa situazione di fatto non poteva e non può che condurre a due esiti, entrambi negativi per l'evoluzione e la crescita delle ricerche sull'uomo. Da una parte si procede ad una lettura errata del potenziale d'informazione latente della carta di distribuzione. Oppure il potenziale d'informazione della carta viene semplicemente tralasciato o trascurato perché non identificabile e fruibile dal ricercatore tramite le metodologie corrette. In base a tutto questo si può giungere ad una considerazione[31]: se è vero quanto detto prima, si potrebbe infatti affermare che la carta geografica, intesa come elemento di sintesi delle distribuzioni spaziali, non è mai effettivamente servita a processi di incremento della conoscenza scientifica sull'uomo.

Il ruolo della cartografia all'interno dei processi di ricerca può essere circoscritto a pure funzioni di tipo illustrativo. Raramente la carta trova posto nelle fasi di formulazione di nuova conoscenza: risulta difficile estrarre da essa qualcosa di nuovo, qualcosa che non si sappia già[32].

Si potrebbe obbiettare però che tutto questo sia semplicemente un'esagerazione. Qualcuno potrebbe giustamente affermare che l'impossibilità di fornire o formulare qualsiasi tipo di considerazione o riflessione su una carta come quella della figura 1.5 dipenda dal fatto che non si dispone di uno sfondo topografico che serva da contesto per potere compiere una lettura di tipo qualitativo; lo sfondo bianco sostanzialmente impedirebbe o limiterebbe ogni modalità di interpretazione o aggancio alla realtà. Inoltre manca ogni sistema di riferimento cartografico come orientamento o scala. Potrebbe essere dunque che il problema vero consista nel fatto che quella della figura 1.5 non sia effettivamente una carta geografica. Inoltre sarebbe necessario sapere cosa rappresentano i simboli utilizzati: cosa sia effettivamente quella rete di distribuzione. In effetti, un gruppo di punti sparpagliato ci dice ben poco sulla natura della rete analizzata: sono case, stelle, città o piante? Illustrare il contesto incrementando i dati a disposizione significherebbe fornire all'osservatore le condizioni basiche dentro le quali si è mosso ad oggi ogni ricercatore che desiderasse affrontare lo studio diretto di una carta geografica.

[30] Uno dei contributi più importanti nell'ambito dell'applicazione della logica spaziale in campo etnoarcheologico è Hodder 1982.

[31] Matthews 1985, pp. 11-13.

[32] Hodder, Orton 1976 pp. 13-14.

La rete di distribuzione della figura 1.5 si riferisce, dunque, alla maglia delle pievi tra le città di Firenze, Siena e Volterra ricavata dalla georeferenziazione delle *Seconde Decime*[33]. I punti in questione sono le pievi della fine del XIII secolo. Lo scopo di questa ispezione resta sempre quello di cogliere qualche elemento che possa fornire delle indicazioni utili alla comprensione della struttura spaziale di tale rete di stanziamento.

Sapere a cosa si riferiscano, o cosa rappresentino effettivamente i simboli della figura 1.5 potrebbe non essere sufficiente per cogliere una tendenza generale. Di fatto, osservare la distribuzione di punti sapendo che si tratta in realtà di chiese battesimali non aiuta molto. Allora, come è stato detto sopra, forse uno sfondo topografico potrebbe fornire degli spunti utili a cogliere ulteriori elementi di riflessione.

La figura 1.6 mostra la stessa rete di distribuzione delle pievi sulla base rappresentante una cartografia sintetica della zona (rilievi e reticolo idrografico), oltre alla scala e alla posizione delle sedi diocesane. Questi nuovi piani cartografici integrati dovrebbero fornire, in base a quanto detto sopra, un supporto utile a compiere, a questo punto, una valutazione migliore della struttura spaziale della rete di punti. Si può procedere perciò a riformulare all'osservatore la richiesta precedente (cioè di descrivere e spiegare la distribuzione) e attendere dunque questa volta una risposta più organica. Purtroppo, osservando la carta 1.6 bisogna ammettere che, nemmeno questa volta, egli sarà in grado di fornire qualche risposta valida, o comunque immune dal soggettivismo implicito cui si accennava sopra.

Anche se può sembrare un controsenso, la maggiore concentrazione di informazioni (cioè la presenza come sfondo della morfologia e del reticolo idrografico) farà sì che la struttura spaziale della rete plebana venga colta dall'occhio dell'osservatore con ancora maggiori difficoltà. Anche se può sembrare incoerente, le carte geografiche tendono a rendere difficoltoso e complicato il processo di lettura e analisi delle reti di distribuzione spaziale nella misura in cui esse vengono arricchite a livello simbolico da nuovi piani d'informazione.

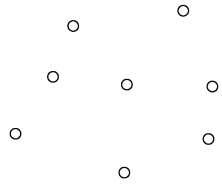
Una prima conclusione o insegnamento di questo esercizio è che risulta alquanto difficile, se non impossibile, compiere delle osservazioni o considerazioni su sistemi distributivi spaziali avvalendosi soltanto del linguaggio comune. Di conseguenza, l'uso della carta geografica tradizionale poco si addice al processo d'analisi spaziale. Come verrà illustrato di seguito, appunto, il più delle volte, i processi d'analisi spaziale avvengono grazie ad un'operazione di astrazione delle entità studiate, in linguaggi di tipo formale-quantitativo.

[33] Giusti, Guidi 1942.

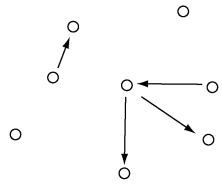
È vero invece che all'interno di questo disordine esiste un *pattern*[34]. Cioè la logica distributiva dei fenomeni rappresentati non è puramente casuale. L'impiego dell'analisi spaziale nello studio delle forme d'insediamento umano si basa sull'ipotesi che l'utilizzo delle maglie di distribuzione degli stanziamenti umani possa restituire informazioni di rilevante valore ed interesse; nozioni che non potrebbero essere in ogni caso recuperate dalle carte di distribuzione con il solo impiego della ragione e dei sensi (la vista appunto). Il motivo per il quale i *patterns* sfuggono all'occhio umano dipende dal fatto che, all'interno di un sistema di distribuzione puntuale come quello delle figure 1.5 o 1.6, la struttura spaziale presenta un numero incredibilmente elevato di rapporti spaziali tra i suoi elementi. È evidente che esistono relazioni punto a punto, ovverosia uno a uno. All'interno della maglia ogni elemento ha comunque un rapporto spaziale (o topologico) con qualsiasi altro elemento. Ma è anche vero che ogni elemento ha un rapporto topologico con tutti gli insiemi che possono essere composti fra i punti rimanenti. A sua volta, ogni insieme di due elementi potrebbe avere relazioni spaziali con tutto il resto degli accorpamenti possibili. Se questo gioco si estende a combinazioni di insiemi di tre, quatto elementi e cosi via (tra l'altro non solo in relazione di contiguità), si può capire come il numero di rapporti topologici possibili aumenti esponenzialmente; anche nel caso di una rete composta da un numero tutto sommato contenuto di punti come quella della figura 1.5. Lo studioso, ad un primo approccio, non può che arrendersi a qualificare e valutare come caotica la grande quantità di possibili rapporti spaziali. Cioè una struttura che apparentemente non possiede alcuna forma d'ordine identificabile o riferibile alla logica umana.

Eppure un ordine all'interno di tale rete di distribuzione esiste certamente. Non solo, come è stato detto precedentemente, questa rete di distribuzione funziona, ma rappresenta, fra le altre cose, anche un sistema. Non solo esiste un *pattern*, ma, come è stato affermato prima, le trame o tendenze distributive, insieme ad altre caratteristiche spaziali, contengono ed esprimono l'essenza stessa caratterizzante un qualsiasi sistema di distribuzione. Così le caratteristiche distributive di una maglia di stanziamento umano rappresentano di per sé un indicatore o un attributo particolare (e spesso unico) delle reti di stanziamento umano. I valori ricavabili dall'approccio quantitativo alle maglie d'insediamento costituiscono un elemento fondamentale perché capace di fornire indicazioni di grande interesse per qualunque processo d'indagine scientifica che abbia come oggetto un fenomeno di distribuzione spaziale.

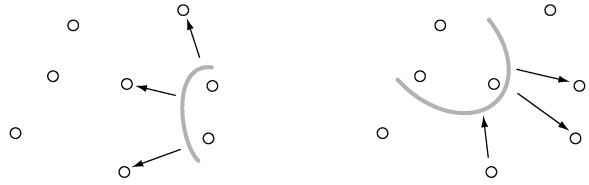
[34] In questo contesto, *pattern* viene utilizzato per indicare «strutture ricorrenti» identificabili all'interno dei sistemi distributivi spaziali. Non esiste una traduzione esatta nella lingua italiana per il termine inglese *pattern*; i termini struttura, trama e forma possono solo in parte rendere il suo vero significato. Vedi Pievani, *Nota del Curatore*, in Eldredge 2000, p. XLI.



esempi di rapporti *1 a 1*



esempi di rapporti *1 a molti*



esempi di rapporti *molti a molti*

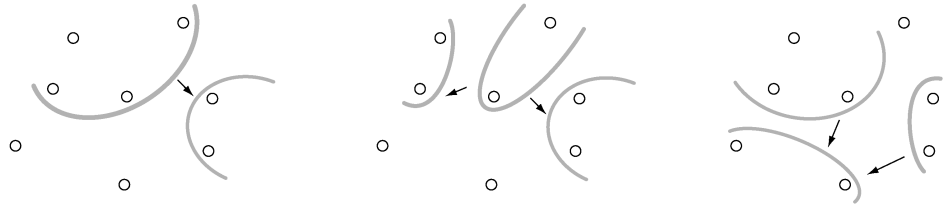


Fig. 1.7. Data una certa maglia di distribuzione puntuale, è possibile stabilire un numero elevatissimo di possibili rapporti topografici.

Proprio per essere un processo di quantificazione formale, il tipo di elementi ricavabile dalla sua applicazione si caratterizza per un grado di obiettività elevato. Questo significa che, se i processi di quantificazione spaziale vengono applicati con sufficiente rigore da due diversi ricercatori, entrambi arriveranno ad un risultato identico o, in ogni caso, simile. Il grado di obbiettività deriva appunto dal fatto che i dati relativi alla distribuzione sotto esame non provengono più da osservazioni relative e parziali ma piuttosto da processi di misurazione assoluta, al di fuori di qualsiasi forma di percezione soggettiva. Naturalmente questa è una valutazione un po' esagerata dato che l'applicazione esatta delle stesse procedure d'analisi è sostanzialmente impossibile. Non esiste un unico modo di misurare. I metodi analitici possono essere migliorati e adattati a condizioni specifiche. Perciò fine dell'approccio quantitativo non dovrebbe essere ottenere risultati identici, quanto fornire dati per i quali possa essere valutato all'interno di un quadro condiviso, il grado di attendibilità.

Se è vero che il tipo di dato prodotto dall'analisi spaziale può essere considerato imparziale, è anche vero che questo permette, all'interno di una cornice più ampia, di giungere al confronto concreto e fondato di due sistemi di distribuzione spaziale. Ad esempio, la comparazione di un'altra sezione territoriale della rete delle pievi con quella illustrata nella figura 1.6. Naturalmente, questo non sarebbe stato neppure pensabile prima, visto che non era fattibile nemmeno giungere alla definizione obiettiva delle caratteristiche distributive di un'unica maglia; figurarsi di due[35].

Di fronte alla possibilità di un'osservazione e di un esame tradizionale impregnato di un soggettivismo implicito, l'analisi spaziale offre la possibilità di usufruire di un rilievo preciso e assoluto. Ma l'aspetto che più conta è che tale rilievo potrà essere sempre confrontato con altre misurazioni.

L'analisi spaziale va ben oltre questi aspetti. La figura 1.8 illustra tre campioni della rete delle pievi alla fine del XIII secolo. Essi coprono tre aree ben distinte e non sovrapposte. Pur essendo parte della stessa maglia, i tre campioni si differenziano a livello geografico. L'analisi spaziale consentirebbe di misurare e valutare, all'interno di un'unica maglia di stanziamento, le differenze interne, consolidatesi lungo il loro processo formativo. Ad esempio, nel caso delle pievi, si può immaginare che la spinta o la motivazione all'origine delle fondazioni sia stata la stessa: quella di portare il fonte battesimale alla popolazione. In questo caso, il processo di quantificazione permetterà di studiare i modi attraverso i quali il territorio e lo spazio hanno modificato e alterato la strutturazione spaziale della maglia. Semplicemente, una domanda

[35] Sull'importanza di giungere in ultima istanza alla comparazione dei risultati e sull'importanza, in termini qualitativi, dei confronti di più aree geografiche, si veda Macchi 2007, pp. 185-200.

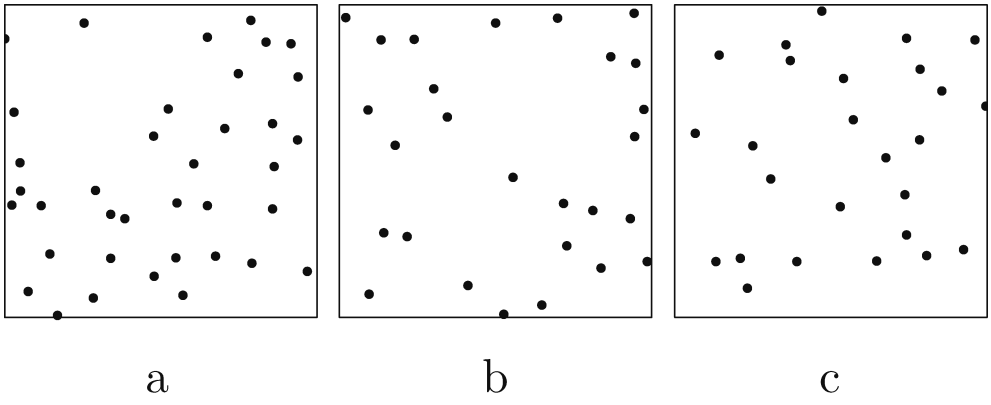


Fig. 1.8. Le reti di distribuzione a, b e c presentano tre campioni diversi della maglia plebana per la Tuscia alla fine del XIII secolo. Con la semplice osservazione e il linguaggio comune non sarebbe possibile giungere all'analisi nè tanto meno al confronto di questi tre frammenti di rete. L'analisi spaziale consentirebbe invece non solo una loro comparazione, ma anche di stabilire degli indici capaci di indicare il grado di affinità o discrepanza al loro interno.

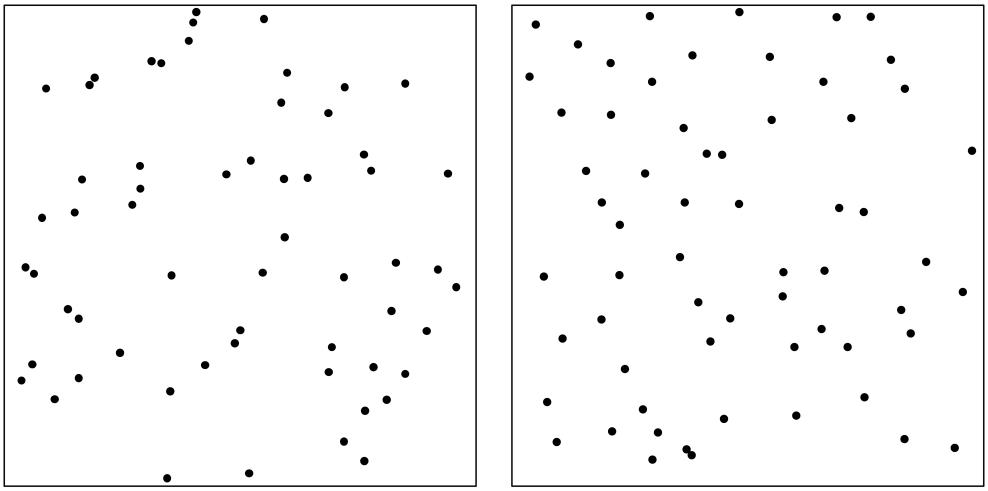
del tipo "se e come i tre campioni si differenziano" sarebbe stata improponibile al di fuori di un approccio quantitativo.

Ad esempio, non è possibile distinguere tra un sistema di distribuzione generato casualmente (*random*) da uno prodotto da un processo caotico, come possono essere le forme di stanziamento umano. Anche se, a occhio nudo, entrambi i sistemi potranno apparire disorganici ed equivalenti, le quantificazioni delle loro caratteristiche a livello spaziale possono presentare grandi differenze.

Per esempio, le distribuzioni puntuali della figura 1.9 possono apparire molto simili. Entrambe presentano un disordine che le accomuna. La prima è stata generata casualmente dal calcolatore, mentre la seconda rappresenta la distribuzione di pievi in un settore della Toscana medievale. L'analisi spaziale può essere considerata come la disciplina che studia, attraverso le particolarità delle distribuzioni, le caratteristiche dei sistemi o processi che le hanno generate.

1.7 Analisi spaziale e ricerca storica

Uno dei settori che potrebbero beneficiare maggiormente della comprensione delle variabili spaziali è proprio la storia degli insediamenti. Basti pensare che una delle principali potenzialità della geografia quantitativa è la possibilità



distribuzione casuale distribuzione caotica

Fig. 1.9. Differenze tra caos e casualità. Agli occhi del ricercatore esistono pochissime differenze tra sistemi distributivi prodotti da processi caotici e sistemi prodotti da processi casuali. L'analisi spaziale può contribuire a distinguere tra questi due tipi di reti. Per esempio, le distribuzioni puntuali della figura possono apparire molto simili: entrambe presentano un disordine che le accomuna. La prima è stata generata casualmente dal calcolatore, mentre la seconda rappresenta la distribuzione di pievi in un settore della Toscana medievale.

di produrre nuove "informazioni"[36] in modo semplice e preciso. Questo fattore assume ancora più valore se inquadrato all'interno del contesto tradizionale della ricerca storica, dove si verifica non raramente una mancanza di fonti e documenti o, in ogni caso, dove ogni informazione costituisce un elemento utile e, si potrebbe aggiungere, attendibile per la formulazione delle sintesi. Gli storici, in particolare i ricercatori impegnati nello studio dei periodi più lontani, non sono costretti soltanto a confrontarsi con gli abissi documentari delle cosiddette età oscure; anche per periodi più ricchi di fonti essi devono fare spesso i conti sia con una grave carenza di documenti, sia con i limiti di alcune di queste. Si prenda ad esempio il caso della storia degli insediamenti, disciplina che più di altre può servirsi dell'eventuale applicazione dell'analisi spaziale nel proprio settore d'indagine. Si può affermare in generale che, per qualsiasi periodo storico, esista un gruppo abbastanza ampio di insediamenti per i quali non si dispone che di qualche vaga e sporadica menzione indiretta.

[36] Per informazione s'intende il risultato di una fase analitica che abbia previsto l'elaborazione di dati di qualsiasi genere. Macchi 2001.

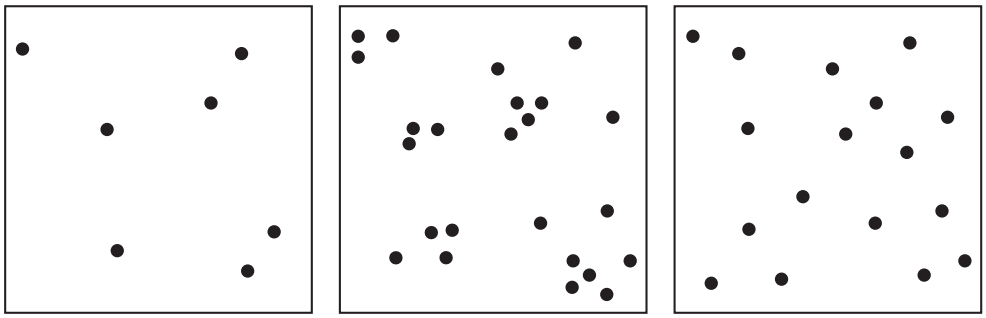


Fig. 1.10. L'applicazione dell'analisi spaziale in campo storico riguarda anche l'evoluzione delle reti di stanziamento umano. Uno dei modelli geografici più importanti che tiene conto della trasformazione nel tempo delle maglie insediative è quello di Hudson che si basa, in sintesi, sulla loro evoluzione nei seguenti tre diversi stadi. *Colonizzazione*: In questa fase gli insediamenti occupano il paesaggio senza seguire nessun ordine specifico. *Crescita*: Gli insediamenti originali danno vita a nuovi stanziamenti determinando una crescita della densità. In questa fase si registra anche un processo di aggregazione. *Competizione*: In quest'ultima fase del modello i livelli di densità determinano processi di selezione naturale che comportano la diminuzione del numero di insediamenti e presumibilmente la definizione di una distribuzione omogenea. Hudson 1969, pp. 365-371.

Senza voler forzare in senso critico il complesso delle informazioni a disposizione degli storici, si potrà concludere che il materiale di cui dispongono è sostanzialmente scarso; certo non si potrà mai giungere all'affermazione che esso sia inadeguato. La storia è stata sempre comunque fatta con le fonti a disposizione. Resta il fatto che l'ambizione dell'acquisizione di nuove informazioni rimane uno degli obiettivi primari di qualsiasi storico. Sia che si tratti dell'opera di spoglio di documentazione scritta o cartografica, che del lavoro di ricognizione o scavo, lo storico ambisce, in ogni caso, all'acquisizione di nuovi saperi utili alla costruzione delle proprie sintesi.

In questo però la storia non si differenzia dal resto delle scienze. La scarsità delle informazioni è semplicemente uno dei maggiori scogli per qualsiasi scienza, sia essa empirica, teorica o sperimentale: la produzione di nuova conoscenza si è sempre scontrata con problematiche di questo tipo, ed è proprio nel superamento di simili difficoltà che essa consiste. Questo processo avviene anche grazie all'evoluzione metodologica degli apparati, sia teorici che empirici, che regolano l'andamento delle ricerche. In questo senso, l'acquisizione di nuovi metodi e la costruzione di nuovi strumenti ha come scopo di base solo

il desiderio di sapere cose nuove. Così esiste, in quello che conosciamo come scienza, un naturale istinto verso l'acquisizione di ulteriori informazioni, senza che però il ricercatore sia effettivamente tenuto a sapere in anticipo la natura e le caratteristiche del nuovo sapere che egli potrà produrre grazie alla sua evoluzione metodologica. Galileo inventa il telescopio perché desidera spingere il suo occhio più in là. Egli però, prima di inventarlo, non poteva conoscere il genere e le caratteristiche delle cose che avrebbe visto.

Il merito principale della geografia quantitativa non è quello di produrre informazioni attendibili; qui riferendomi a un'obiettività derivante dalla sua natura quantitativa o numerica. L'aspetto più importante e interessante è la costruzione di nuove informazioni. L'estrazione dei *patterns* dal caos apparente delle carte di distribuzione. Naturalmente, nell'indicare le possibilità di un percorso verso la comprensione formale dello spazio, si vuole segnalare che tale opportunità costituisce per certi versi una strada fondamentale nella acquisizione organica e concreta di gran parte dei fenomeni che rientrano all'interno della fenomenologia sociale.

2
Modelli, modelli geografici e modelli scientifici

In senso generale, si potrebbe affermare che scopo della scienza sia, nell'interesse della specie umana, la comprensione della realtà. Anche se questa resta una valutazione riduttiva, contiene gran parte del significato profondo del rapporto che esiste tra specie umana e quella che comunemente viene conosciuta come scienza e il suo prodotto principale: il sapere o la conoscenza. Certamente sussistono divergenze sia nella storia che nella filosofia della scienza sui fini ultimi di questa comprensione della realtà [1]. Resta il fatto però che, la conoscenza attraverso la ricerca scientifica dipende esclusivamente dalla capacità di comprendere e studiare l'universo. In altre parole, di essere in grado di leggere e capire la realtà.

Naturalmente vi sono diverse forme e modi attraverso i quali la scienza giunge alla produzione di nuova conoscenza. Tra i più diffusi – più di quanto non si pensi – ed efficaci vi è quello che solitamente viene definito come "modello scientifico" o, più semplicemente, "modello". Si potrebbe affermare, senza paura di esagerare, che di fatto esso costituisce per tale scopo, lo strumento più adatto per giungere cioè alla produzione di nuova conoscenza[2].

Questo capitolo ha come scopo quello di compiere una breve riflessione sul concetto di modello e sui diversi significati che questo termine può ricoprire o acquistare all'interno della dialettica della geografia e dei processi di studio dei paesaggi umani. Come verrà illustrato di seguito, i modelli sono infatti un elemento caratterizzante della geografia; basti pensare, ad esempio, che ogni singola carta geografica rappresenta un modello.

Il Dizionario di Geografia Umana di Johnston, Gregory e Smith definisice il modello come: «An idealized and structured representation of the real[...]

[1] Lakatos 1995.

[2] Sull'utilizzo della metafora di "strumento" come definizione del "modello scientifico" fare riferimento a Cartwright, Shomar, Suárez 1995, pp. 137-149; Fox Keller 2000, pp. 72-86.

Model-building [...] depends on the so called analogue theory, which regards models as selective approximations which, by the elimination of incidental detail or noise, allow some fundamental, relevant or interesting aspects of the real world to appear in some generalized form.»[3] Questa breve definizione sintetizza efficacemente tutti gli elementi chiave per la comprensione della natura dei modelli scientifici. Ma, come verrà chiarito in seguito, il concetto di modello è molto esteso e ricco di significati; richiede dunque una trattazione precisa e dettagliata. Non sono pochi i fraintendimenti che si possono generare a causa dello spettro di significati, valori e utilizzi dei diversi tipi di modelli. Sarà scopo delle prossime pagine introdurre ad alcuni aspetti e significati di questa parola.

La volontà di chiarire le differenze e il valore dei possibili significati del termine in modo preventivo è dovuto anche al fatto che il concetto di modello, nel campo della geografia, oltre ad un significato teoretico, ne possiede anche uno metodologico e perfino uno tecnico. In pratica, nel discorso geografico gli si conferisce più significati; e spesso con variazioni minime. Da qui la necessità almeno di tentare di chiarirli con precisione, oltre che descrivere i contesti nei quali se ne fa uso.

2.1 La natura dei modelli scientifici

È un mondo complesso quello in cui la specie umana abita. E la complessità risiede anche nell'immensa quantità di rapporti che caratterizzano appunto quella che noi uomini conosciamo come realtà[4]. È altrettanto vero, però, che la ricerca scientifica non potrebbe, per sua stessa natura, arrendersi al caos che la circonda. Con il concetto di caos si fa riferimento non tanto all'idea di disordine o confusione, quanto ad un ordine superiore tipico dei sistemi complessi. Proprio per questo motivo, fa parte dei processi di ricerca lo sviluppo e l'identificazione di strumenti, anche concettuali, sempre più sofisticati per risolvere e semplificare un ostacolo (quello della complessità appunto) che indurrebbe senza alcun dubbio al rischio di far interrompere qualsiasi processo o forma d'indagine scientifica[5].

[3] Johnston, Gregory, Smith 1986, p. 301.

[4] McDermott, O'Connor 2003.

[5] Per un'esposizione organica sul tema della complessità in geografia fare riferimento a Malanson 1999: «As geographers studying place, we have a tendency to include as many details as possible in simulation models. The goal of the science of complexity, [...], should be considered useful in helping us to represent place as simply as possible and enabling us to find parsimonious models for the emergence of places. These models will not predict the behaviors of particular systems, but should address types of patterns and variability within which reductionism will remain necessary.»

Lo studio di un mondo così ampio e caratterizzato da rapporti tanto complessi, richiede dunque strumenti efficaci ai fini della semplificazione dei processi cognitivi. La chiarificazione di una rete così contorta di rapporti appare perciò come una necessità. Ed è per questo motivo che, dagli inizi della ricerca scientifica, e probabilmente anche da prima, lo strumento conoscitivo più importante sviluppato dalla specie umana è stato il modello.

Gran parte della conoscenza (qui nel senso di semplice sapere) si fonda su modelli, intesi come teorie sulla realtà. Una dimostrazione chiara di questo concetto è il paradosso di Hempel[6]. Una delle sue interpretazioni ci permette di capire per quale motivo gran parte del sapere umano si basa più su modelli che non su certezze. Dunque, è del tutto lecito affermare che *tutti i corvi sono neri*. Un'affermazione così ovvia e semplice però, non può essere considerata vera in termini assoluti. Di fatto, per essere vera, un gruppo di osservatori avrebbe dovuto avere la possibilità materiale di verificare il colore di ogni singolo esemplare di corvo. Anche di quelli che vivono nelle aree più remote della terra. E anche così non sarebbe sufficiente dato che anche qualora fosse possibile verificare il colore di tutti corvi viventi, sarebbe necessario controllare il colore dei corvi già scomparsi, e naturalmente anche di quelli futuri. *Tutti i corvi sono neri* non è altro che una teoria. Ogni qual volta viene osservato da qualcuno un nuovo esemplare di corvo nero, la teoria di partenza viene confermata e la sua validità sarà da considerare ancor più plausibile. Nonostante ciò, in termini puramente epistemologici, è impossibile accettare come vera al 100% l'affermazione *tutti i corvi sono neri*.

Naturalmente questa idea di modello come "teoria probabile" (o poco probabile) va estesa ulteriormente e resa più organica nel campo della ricerca scientifica. Bisogna ricordare sempre che i modelli possono essere assimilati a processi di traduzione della realtà in schemi teorici contraddistinti da un certo livello di probabilità.

I modelli scientifici (o i modelli in generale) sono, nella quasi totalità dei casi, caratterizzati dall'essere sintetici e significanti. La realtà è così riportata ad uno scheletro semplificato capace in ogni modo di contenere e descrivere gli elementi e i rapporti più rappresentativi ed importanti secondo l'ottica e l'interpretazione di colui che costruisce il modello.

Col termine "sintetico" si intende che il modello costituisce, in ogni caso, una semplificazione del mondo. Esso prevede perciò l'eliminazione di ampie porzioni della realtà. Solo un piccolo segmento ne può essere preso in considerazione. Ed è per questo che, per nessun motivo e in nessun momento, il modello di ricerca scientifico potrebbe essere considerato come possibile du-

[6] Hempel 1943 e Hempel 1946.

plicato o come riproduzione esatta della realtà. La strategia fondamentale e l'efficacia del modello scientifico sta proprio nella sua capacità di semplificare.

Nel caso si procedesse alla costruzione di un modello molto ampio che prevedesse la quasi totalità, o addirittura la totalità, dei rapporti e fattori percepibili dal ricercatore, il modello andrebbe ad assumere l'essenza e la natura di copia o duplicato astratto della realtà. Questo significherebbe riprodurre una copia esatta della complessità implicita nella realtà. Per cui il ruolo principale del modello, cioè quello di semplificarla in termini astratti, verrebbe immediatamente meno. La natura e lo scopo di un modello utile, anche a costo di falsare la realtà, sono sempre quelli di semplificare.

A questo proposito pagine e pagine sono state scritte da filosofi della scienza. È stato però Jorge Luis Borges ad avere esposto efficacemente e con maggiore chiarezza la portata della questione nel suo poema *Del rigor de la ciencia*, peraltro intimamente legato con la disciplina geografica:[7]

> «. . . En aquel imperio, el Arte de la Cartografía logró tal Perfección que el mapa de una sola Provincia ocupaba toda una Ciudad, y el mapa del Imperio, toda una Provincia. Con el tiempo, esos Mapas Desmesurados no satisficieron y los Colegios de Cartógrafos levantaron un Mapa del Imperio, que tenía el tamaño del Imperio y coincidía puntualmente con él. Menos Adictas al Estudio de la Cartografía, las Generaciones Siguientes entendieron que ese dilatado Mapa era Inútil y no sin Impiedad lo entregaron a las Inclemencias del Sol y de los Inviernos. En los desiertos del Oeste perduran despedazadas Ruinas del Mapa, habitadas por Animales y por Mendigos; en todo el País no hay otra reliquia de las Disciplinas Geográficas.»

Il secondo punto sul quale riflettere è che il modello scientifico prende in considerazione solo quegli oggetti e quelle relazioni che il ricercatore considera importanti ai fini della comprensione e dell'illustrazione dei fenomeni indagati[8]. Per questo motivo i modelli tendono ad essere non solo sintetici,

[7] In quell'impero, l'arte della cartografia raggiunse tale perfezione che la mappa di una sola provincia occupava tutta una città, e la mappa dell'impero, tutta la provincia. Con il passare del tempo, quelle mappe sconfinate non soddisfecero più e le scuole di cartografi rilevarono una mappa dell'impero che aveva la grandezza dell'impero e coincideva perfettamente con questo. Meno assoggettate allo studio della cartografia, le generazioni seguenti capirono che questa estesa mappa era inutile e non senza crudeltà la consegnarono alle inclemenze del sole e degli inverni. Nei deserti dell'ovest resistono a brandelli rovine della mappa, abitate da animali e mendicanti; in tutta la nazione non ci sono altre reliquie delle discipline geografiche.

[8] Popper 1995b, pp. 218-219; Sharov 1992.

ma soprattutto significanti. Nel senso che mettono in risalto quegli elementi e quelle dinamiche che sempre secondo chi li studia, hanno un significato rilevante. Nel farlo naturalmente egli compie una scelta: il significato del modello è sempre da mettere in rapporto all'interpretazione compiuta dal ricercatore.

Come prodotto specifico di settori e aree disciplinari, i modelli intesi come schemi significanti al di fuori di una logica interdisciplinare, possono essere limitanti. Vale a dire che, come prodotto di un'unica disciplina, tendono a mettere in evidenza o a conferire significato solo a quei fattori o elementi che quella disciplina tratta. Così l'antropologia sviluppa modelli culturali, la sociologia modelli sociali, l'economia economici, la geografia spaziali ecc. L'importanza dell'interdisciplinarità sta proprio nella capacità di integrare in teorie comuni una visione plurale della realtà.

2.1.1 La costruzione dei modelli

Il processo di formulazione di un modello può essere inteso come la definizione di una serie finita di idee e concetti sulla realtà. Riguarda in particolare la sostituzione di una o più parti della realtà con una serie di rappresentazioni simboliche a livello astratto grazie all'utilizzo di forme di linguaggio specifiche[9]. Si potrebbe dire però che ogni modello non è altro che la definizione di uno schema teso, attraverso simboli e regole, alla simulazione parziale o totale della complessità di una determinata porzione della realtà[10]; nel caso del modello scientifico, di quella parte sotto indagine.

I modelli hanno un intrinseco parallelismo o analogia con i "sistemi". Di fatto, l'essere umano ha da tempo sviluppato una visione sistemica della realtà. Per l'intelletto umano i sistemi sono ovunque ed è questa la chiave attraverso la quale ci siamo abituati ad osservare il mondo.

Con il termine modello spesso ci si riferisce all'enunciazione di un sistema. Occorre ricordare che il concetto di sistema, e la sua influenza nella formulazione dei modelli, ha subito un'evoluzione di non poco conto. Dall'antichità greca, quando i sistemi erano concepiti come entità organiche, al Rinascimento, con strutture prevalentemente meccaniche – frutto anche delle innovazioni tecniche, la rivoluzione newtoniana e quelle industriali successive – per ritornare, nel XX secolo, alla riformulazioni di modelli organici di tipo olistico.

I modelli vengono prodotti in diverse forme e tipologie. Essi possono descrivere gli elementi attraverso equazioni, schemi o algoritmi di flusso, ricostruzioni tridimensionali al computer, grafici statistici o semplicemente a parole,

[9] Ackermann 1966, p. 314-315.
[10] Meadows 1957, p. 4.

con il linguaggio corrente[11]. Di fatto, il modello non è tanto legato al metodo d'illustrazione o descrizione, quanto al processo d'astrazione che compie come veniva detto prima. In linea di principio, qualsiasi metodo illustrativo possa soddisfare le necessità di esposizione può essere usato liberamente per trasmettere il modello[12] .

Un modello costituisce perciò – e viene sottolineato ancora una volta – uno schema teorico relativo a una porzione specifica della realtà. È opportuno ricordare che la scelta stessa delle parti e delle relazioni non riflette mai la realtà quanto, invece, l'interpretazione che il ricercatore, in modo del tutto soggettivo, fa di essa. Queste implicazioni hanno un peso ancora maggiore se si pensa al fatto che il ricercatore utilizza il modello, appunto, per conoscere fenomeni o elementi che costituiscono o fanno parte del modello stesso. Il che significa che esso è composto da elementi e parti che comunque sfuggono alla comprensione stessa del ricercatore. Se non fosse così, non avrebbe molto senso studiare questi fenomeni, né tanto meno sforzarsi nella costruzione di un modello ad essi riferito. Ammesso che vi siano settori della realtà e dell'universo interamente compresi dall'uomo, non vi sarebbe motivo di continuare a studiarli. Il modello non è la fine di un percorso d'indagine, bensì la base congetturale di partenza.

Questa ultima considerazione porta ad una conclusione molto importante. Non solo i modelli di ricerca scientifica sono incompleti perché sintetici, ma, per loro stessa essenza e scopo, essi sono anche imperfetti. Appartengono in ogni caso a stadi intermedi della conoscenza. E, considerando il fatto che i processi di ricerca non possono, proprio per la loro natura, essere considerati sotto nessun aspetto definitivi, i modelli stessi non potranno essere portatori del livello di compiutezza o perfezione che spesso si pensa essi possiedano. I modelli sono per definizione imperfetti.

Essi cambiano e si evolvono costantemente. Dato che in sostanza essi non sono altro che una falsificazione simbolica della realtà, va da sé che esisterà sempre un modo migliore per renderla in modo più preciso e sofisticato; questo cercando di mantenere il modello il più sintetico e semplice possibile. In larga misura, tale processo dipende anche dall'evoluzione della società stessa che lo produce: la sua consapevolezza sul mondo, la sua capacità di osservar-

[11] Wegener 2000.

[12] Wilson 1972, p. 32: «A theory is a set of propositions which purports to explain the structure of some system and/or how the system develops. That is, we may be interested in patterns or process or, ideally, both. The key term in this definition is 'explain'. Note that a theory does not have to be quantitative or mathematical by definition. The proposition may be expressed in a variety of languages. It so happens that, in the case of some complicated systems, explanation can only be adequately achieved with the help of mathematical language».

lo, l'introduzione di nuovi strumenti o metodi, ma soprattutto la capacità di formulare concetti logici grazie all'uso del linguaggio, anche di tipo formale. Non esistono modelli statici (o chiusi) nel tempo.

Un esempio chiaro di questo fatto è, per l'appunto, quello che viene considerato come il primo modello scientifico: quello del sistema solare[13]. La sua meccanica, considerata dalla maggior parte dell'umanità come dato acquisito, è stato frutto di una costante evoluzione ancor oggi non conclusa. Il modello relativo al sistema solare si è così evoluto nelle seguenti fasi:

- Nell'antichità diverse sono le ipotesi sulla natura della volta celeste e della organizzazione degli astri nel cielo. Queste si intrecciano con elementi mitico-religiosi che servono come impalcatura della logica causa-effetto e della struttura dinamica.
- Nel II secolo d.C. Tolomeo introduce un modello che vede la Terra al centro dell'universo[14]. Questo è costituito da un'immensa sfera puntellata di stelle al suo interno. Dentro questa sono presenti altre sfere concentriche e trasparenti che sostengono i pianeti del sistema solare che vi sono incastonati. L'intero meccanismo si regge grazie alla posizione di puntello centrale della Terra al centro dell'universo. Nell'*Almagesto* Tolomeo dedica ampio spazio ad una serie di computi che permettono di calcolare e prevedere il moto degli astri più importanti. In particolare il Sole e la Luna.
- Per molti secoli il modello di Tolomeo rimane il migliore ed unico modo per spiegare la meccanica celeste. Durante i primi anni del XVI secolo, grazie alle sue osservazioni, Copernico introduce il concetto di centralità del Sole. È la Terra dunque a girare attorno al Sole. Il suo modello prevede ancora l'esistenza degli emicicli; oltre alla circolarità delle orbite disegnate dai pianeti attorno al Sole.
- Poco dopo, Tycho Brahe, grazie all'osservazione del movimento delle comete, respinge l'idea dell'esistenza di emicicli o elementi di sostegno simili, ma compie anche un passo indietro collocando nuovamente la Terra al centro dell'universo. Nel suo modello sono solo gli altri pianeti (Mercurio, Venere, Marte, Giove e Saturno) a muoversi attorno al Sole.
- Da lì a poco, Johan Kepler (1571-1630) reintroduce definitivamente la posizione centrale del Sole e giunge ad una perfetta comprensione del moto dei pianeti con l'ipotesi della forma ellittica delle loro orbite.
- È poi nel XVII secolo, grazie ad Isaac Newton e alle tre leggi sul principio della dinamica, che si può giungere a formulare un modello "accettabile" per descrivere la struttura e la dinamica del sistema solare. Con i

[13] Popper 1995b, pp. 220-222.
[14] Toomer 1998.

principi da lui introdotti diviene possibile prevedere per un determinato momento, e con grande precisione, la posizione esatta di ogni pianeta.

- Poco dopo Laplace (1749-1827) pubblica la *Meccanica Celeste*. Il testo dipinge un universo – basato sulla logica newtoniana – molto simile ad un immenso ingranaggio celeste contemplato da Dio; dove ogni stella o pianeta trova la sua posizione in un determinato luogo e momento.
- Nonostante il grande contributo della meccanica newtoniana, durante i secoli XVIII e XIX, ben presto ci si accorse che non è possibile, come invece lo è per gli altri pianeti, giungere ad una previsione assoluta relativa al moto di Mercurio. Questo errore, associato allo sconforto degli altri scienziati, fa sì che per lungo tempo si ipotizzi l'esistenza di altri pianeti minori e invisibili responsabili dell'errore dei calcoli.
- Successivamente si pone un'altra questione (che porterà poi alla nascita della teoria della relatività). Einstein giunge all'ipotesi che niente sia più veloce della luce. Il che porta dunque a chiedersi cosa sia la gravità. Se niente può superare la velocità della luce allora nemmeno la gravità può farlo. In estrema sintesi, a quel tempo si pensava che, se per ipotesi il sole fosse scomparso, automaticamente la Terra si sarebbe sganciata dalla sua orbita partendo per la sua tangente. Ma, se la luce impiega 8 minuti ad arrivare alla Terra, cosa succederebbe in quell'intervallo? La teoria di Einstein (sulla velocità della luce) e la meccanica di Newton risultano incompatibili. Einstein risolve tutto con l'introduzione di una teoria della relatività e, per il caso del sistema solare, della curvatura spazio-temporale. È solo con l'introduzione delle teorie della relatività di Einstein che finalmente si comprende come lo spazio e il tempo possano subire, sotto certe condizioni, delle anomalie e delle distorsioni[15]. Eventualità plausibili nelle immediate vicinanze del Sole, data la sua massa. L'imprecisione nel calcolo su Mercurio dipende solo dalla distorsione provocata dal Sole stesso.
- Oggi, grazie alla meccanica quantistica, si possiede una visione completamente distinta dell'universo. Le certezze sulla natura del sistema solare e dell'universo appaiono ora più che mai lontane.

Da questo breve *excursus* possono essere tratti almeno due insegnamenti fondamentali. Il primo è che, grazie all'osservazione della realtà, i modelli costituiscono un modo concreto di accrescere il sapere attraverso la definizione di una sua rappresentazione più plausibile (nel senso di probabile). Nonostante la loro semplicità implicita, i modelli costituiscono l'unico mezzo capace di descrivere e aiutare a capire e decifrare la complessità del mondo. Nella

[15] Clemence 1949.

maggior parte dei casi, nell'ambito dell'indagine scientifica moderna, essi rappresentano l'unico strumento a disposizione per avvicinarsi alla comprensione della realtà. Senza di essi, l'efficacia della ricerca sarebbe molto limitata.

La seconda considerazione costituisce un punto ancora più importante. Nella ricerca scientifica non si giunge mai a valutazioni del tipo: *questo modello è corretto* o *questo modello è esatto* o *vero*. Proprio per essere uno stadio intermedio del processo d'indagine, è giusto invece affermare *questo modello è accettabile* o *questo modello è efficace*[16].

La plausibilità del modello è legata a tre elementi fondamentali. In primo luogo un modello deve essere capace di dare una spiegazione ai fenomeni o agli oggetti studiati. Non solo, il modello deve anche essere in grado di fare delle previsioni sugli eventi contemplati o sull'evoluzione dei fenomeni studiati. Infine, il modello deve essere compatibile con l'intero quadro della conoscenza prodotta[17]. Se questi tre aspetti vengono rispettati, si può giungere ad affermare che esso sia accettabile.

È importante ricordare come il motivo per il quale la ricerca astronomica si sia evoluta precocemente nello sviluppo di modelli sia dipeso del fatto che lo studio o l'osservazione della volta celeste ha a che fare con fenomeni che si ripetono periodicamente: il giorno e la notte, il movimento delle stelle o i solstizi sono solo alcuni esempi. Per comprendere i movimenti ed il susseguirsi degli eventi nel cielo, ma soprattutto per giungere a predizioni sufficientemente valide e rigorose sul futuro, si rendeva necessario plasmare uno schema sulla meccanica celeste.

Un fattore importante è appunto che tutti i modelli sopra descritti erano ampiamente diffusi e accettati fino alla nascita di un nuovo modello che, secondo la logica dei paradigmi di Kuhn, doveva spiegare, oltre agli errori e ai vuoti del modello precedente, anche per quale motivo questo sembrasse funzionare. In questo senso i modelli tendono a sviluppare una tradizione di parentela dove un modello può dare vita a uno o più successori[18]. Sarà poi un meccanismo basato su una logica di selezione naturale a stabilire quale sarà il successore più valido. La storia del modello di Tyco Brahe dimostra inoltre come i paradigmi non progrediscano linearmente ma possano compiere passi indietro qualora in quel momento essi vengano considerati come la spiegazione più conforme o plausibile. In definitiva si tratta sempre di una scelta arbitraria compiuta da un gruppo sociale.

[16] Popper 1995b.
[17] Vedi la «priorità dei paradigmi» in Kuhn 1969.
[18] Vedi appunto Ticho Brahe e Keplero.

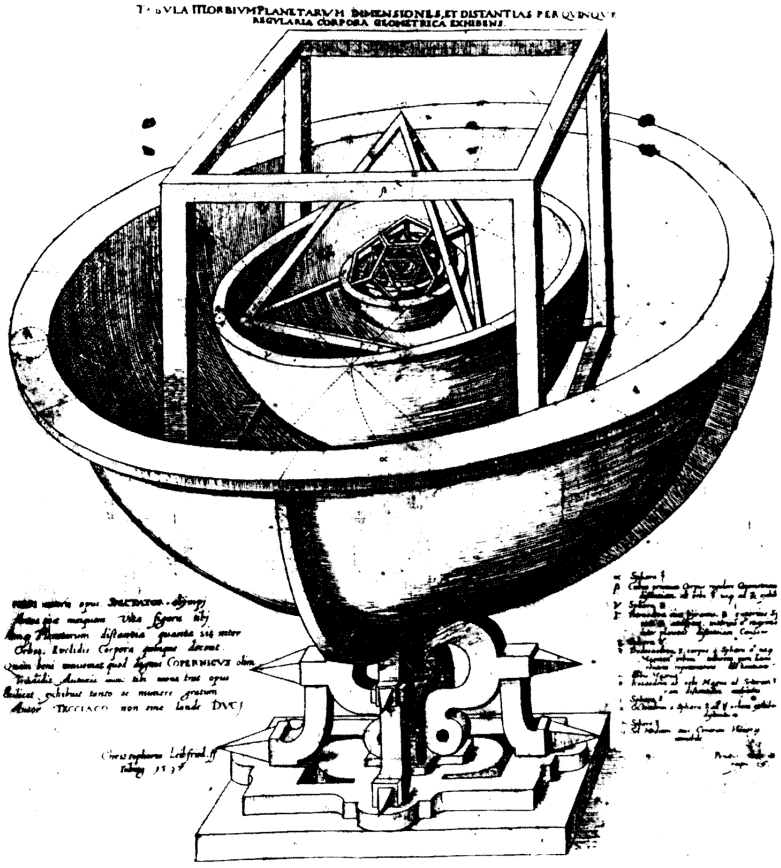


Fig. 2.1. Rappresentazione della sfera armillare di Keplero (*Mysterium Cosmographicum* 1596). Keplero è stato indubbiamente una delle figure chiave nell'evoluzione del modello del sistema solare. È infatti grazie alle sue osservazioni che questo modello, benchè ancora lontano dal fornire previsioni precise sul moto dei pianeti, ha portato alla definizione della struttura generale del sistema solare.

2.2 Modelli geografici

La geografia nel suo complesso è fra quelle discipline che maggiormente hanno sviluppato e fanno correntemente uso di modelli come strumento d'indagine; spesso anche in modo inconsapevole. Di fatto, la carta geografica, l'elaborato che contraddistingue maggiormente il sapere geografico, può – e in molti casi deve – essere considerato come un modello. Le carte geografiche, naturalmente, seguono l'iter evolutivo di qualsiasi modello scientifico.

Oltre alla cartografia, la tradizione della disciplina geografica ha visto la formazione di due filoni o approcci intimamente legati con produzione ed elaborazione di schemi teorici: l'analisi locazionale[19] e la geografia quantitativa[20]. Spesso queste due tradizioni vengono confuse e non distinte nel modo più appropriato. Nell'ambito delle scienze sociali e delle discipline umane, molti pensano che siano di fatto la stessa cosa, quando, in realtà, le due tradizioni, pur essendo entrambe filoni o branche della geografia, sono portatrici di valori e approcci alla conoscenza marcatamente distinti.

Di fatto la tradizione dell'analisi locazionale si rifá ad uno schema di tipo deterministico[21], mentre invece l'analisi spaziale, per sua stessa essenza, si fonda su principi propri della teoria delle probabilità e dunque è più vicina al possibilismo geografico. L'analisi locazionale si basa sulla definizione di schemi nomotetici, cioè mira alla spiegazione della logica delle posizioni dei diversi fenomeni sociali (economia, attività produttive, popolamento e insediamento) nello spazio. Si basa prevalentemente su una logica deduttiva. L'analisi spaziale, invece, mira, benché a un livello formale, ad una descrizione idiografica della realtà, cerca cioè di formulare una serie di modelli probabilistici seguendo un ragionamento induttivo.

In base a queste considerazioni iniziali, si può di conseguenza affermare che nella geografia i modelli possono essere organizzati in tre categorie specifiche. Questa classificazione esprime naturalmente diversi gradi di formalizzazione. I modelli possono dunque essere classificati in[22]:

- cartografici o scalari
- locazionali o deterministici
- probabilistici o statistici[23]

[19] Nome diffuso grazie al volume *Locational analysis* di Hagget 1965. Tradizione definita anche «teorie normative» in Vagaggini 1982.

[20] Qui intesa come analisi spaziale e applicazione di altri metodi statistici.

[21] Qui il termine deterministico è usato in senso logico-spaziale e non riferito al determinismo ambientale di Ratzel e Churcill Semple.

[22] Wegener 2000, pp. 5-6.

[23] Questi modelli verranno trattati nel capitolo 4.

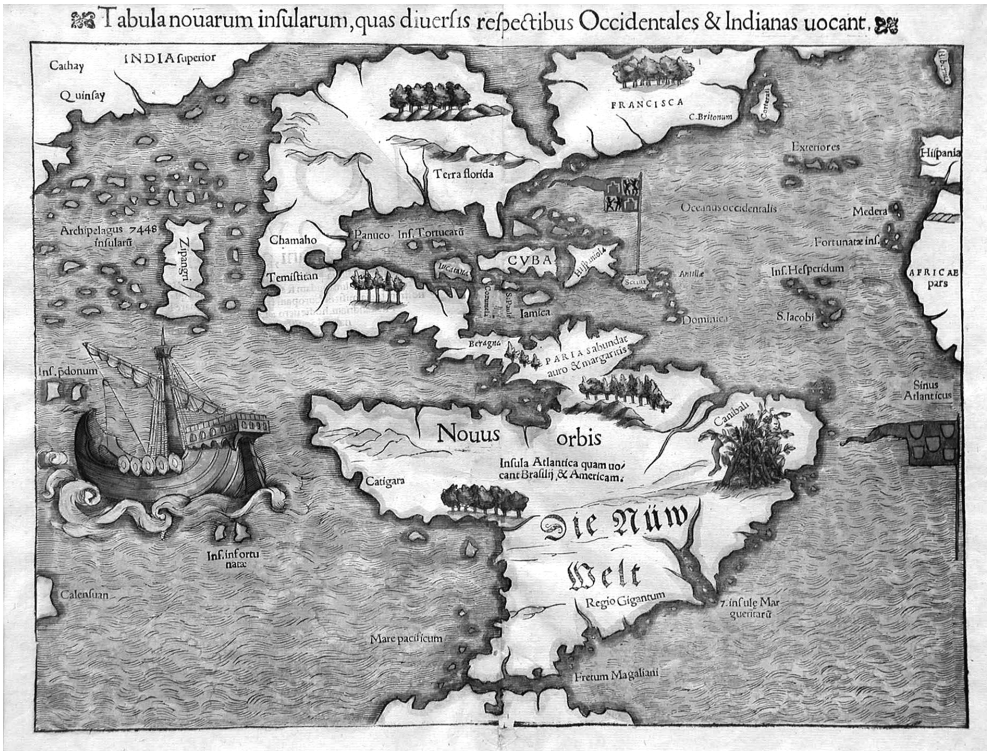


Fig. 2.2. S. Münster, *La Cosmographie universelle*, 1552.

Prima di entrare nel merito di queste tre categorie, occorre segnalare che, se gran parte della logica metodologica nello studio dello spazio riguarda la costruzione di modelli appartenenti a questa classificazione, va detto che frequentemente i modelli prodotti sono anche possibili integrazioni in più forme. L'analisi spaziale può servire, ad esempio, alla costruzione di migliori apparati cartografici. In questo senso si giunge alla produzione di modelli sì di tipo cartografico, ma che non sono altro che la rappresentazione simbolica di modelli probabilistici. In altre parole i modelli possono essere semplificazioni o astrazioni di altri. Nello stesso modo, i modelli locazionali possono essere utili alla definizione di nuove strategie di quantificazione. Ad esempio, un modello locazionale (ed il suo scheletro teoretico di tipo deduttivo) può essere sfruttato come base per definire una migliore strategia d'indagine e misurazione di tipo analitico. Come si può capire, l'applicazione di metodi formali nello studio dei paesaggi umani prevede non tanto l'applicazione esclusiva di questi percorsi, quanto piuttosto la loro integrazione.

2.3 Carte come modelli

Se si prendono in considerazione le carte geografiche, si può notare che queste presentano delle caratteristiche e rispettano le due condizioni fondamentali di qualsiasi modello: sintesi e significatività[24]. Una mappa è sintetica perché appunto, all'interno del foglio di carta, solo una parte delle caratteristiche e degli attributi della realtà può essere rappresentata. La scala e i limiti fisici del supporto non consentono di rappresentare ogni singolo dettaglio. Allo stesso tempo, però, gli oggetti rappresentati sono solo quelli che il cartografo considera fondamentali ai fini della costruzione della mappa.

Oltre a questo, va tenuto conto che le carte geografiche si sono da sempre evolute nel tempo. Le rappresentazioni più dettagliate e precise sono sempre andate parallelamente alla crescita delle conoscenze sull'oggetto geografico rappresentato. Nello stesso modo, le nuove carte sono servite come base di partenza per la costruzione di quelle successive. Sulla base di conoscenze pregresse, le nuove carte vanno sempre a migliorare le precedenti. In questo senso l'evoluzione della cartografia non differisce dalla crescita di qualsiasi altro modello scientifico.

Un caso molto chiaro di questo processo è, ad esempio, la cartografia del Nuovo Mondo. Dalla sua scoperta si può cogliere chiaramente come, grazie alla navigazione e ai viaggi, si sia progressivamente delineato un apparato cartografico più chiaro e dettagliato. In questo processo ebbe grande importanza la tradizione, la copia e la correzione di altre mappe.

Bisogna fare attenzione che, in questa logica, anche la riproduzione con metodi di telerilevamento della superficie terrestre costituisce o deve essere considerata come modello[25]. Anche ortofotocarte o fotografie aeree non possono essere ritenute copie della realtà. La figura 2.3, ad esempio, illustra una porzione della ortofotocarta dell'area intorno al centro di Siena. L'immagine, benché di grande qualità, rappresenta solo una piccola porzione della geografia della zona; peraltro in modo arbitrario, dato che le immagini tradizionali non possono apportare elementi di discrimine e dunque adempiere al principio di significatività del modello. Essa non presenta o non rende esplicite, ad esempio, informazioni relative alla densità della popolazione, al sistema di fognature, alle linee di trasporto pubblico, all'acquedotto, al flusso di pedoni ad una determinata ora del giorno, al prezzo medio al metro quadro degli immobili, ecc. La carta geografica, proprio per essere rappresentazione della realtà, si costituisce necessariamente come schema sintetico e significante di essa. «La cartografia anche quella di oggi – vale a dire la produzione scientifica

[24] Hagget 1997, pp.24-25.
[25] Hagget, 1997, p. 25.

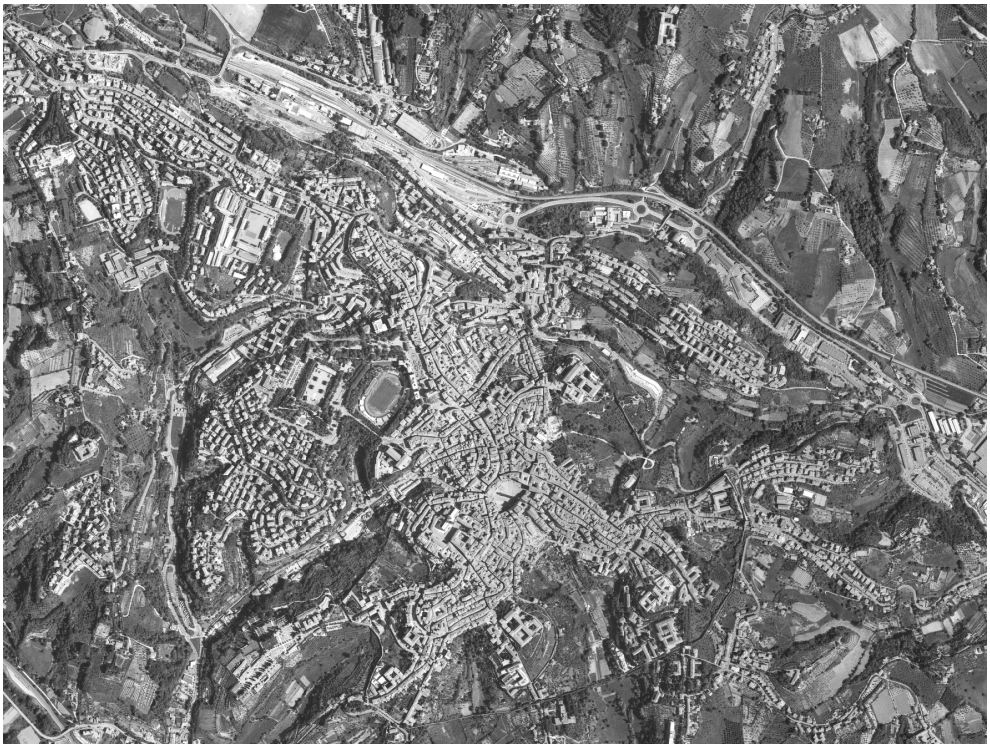


Fig. 2.3. Anche le immagini e la cartografia prodotta con metodi di telerilevamento sono e devono essere considerate modelli. Questa ortofotocarta mostra, in ogni dettaglio, il centro della città di Siena. Non si tratta però di una copia della realtà: l'immagine non mostra infatti elementi quali densità della popolazione, linee di collegamento, prezi degli immobili ecc.

fatta con le foto aeree e con i satelliti – non è mai una vera rappresentazione fedele e integrale del territorio. La cartografia [...] è sempre una costruzione soggettiva [...]»[26].

Nella carta non si può rappresentare tutto; solitamente una mappa illustra uno o, in ogni caso, pochi aspetti della realtà. Essa costituisce sempre una sintesi. Anche l'oggetto stesso viene abbozzato e riepilogato a livello cartografico. A questo si aggiunga anche il problema della scala. Perfino nella rappresentazione cartografica 1 : 25.000 non è possibile rappresentare ogni curva di una strada.

[26] Rombai 2002, p. 31.

Inoltre il cartografo compie, a seconda della scala, una scelta. E in questo senso diviene un documento significativo che illustra quegli aspetti che lui, per i più svariati fini o interessi, desidera mettere in evidenza.

2.4 Modello dei dati e cartografia digitale

Nell'ambito della geografia, e più in particolare in quello della cartografia digitale e i sistemi informativi geografici, il concetto di modello ha acquistato negli ultimi anni un ulteriore significato.

I *Geographic Information Systems* o GIS si basano sull'utilizzo di un cosiddetto "modello dei dati". Questo corrisponde allo schema d'astrazione attraverso il quale il cartografo procede alla rappresentazione grafica della realtà all'interno di una dimensione digitale. In altri termini, il modello dei dati risponde alle necessità di costruzione del documento cartografico con alcune strategie di rappresentazione digitale dello spazio. I modelli dei dati costituiscono cioè parte integrante della "rappresentazione dell'informazione"[27].

Lungo la storia dei GIS sono sostanzialmente due i modelli che hanno raggiunto un maggiore grado di diffusione e che si sono affermati: quello raster e quello vettoriale. Oltre a questi, ci sono anche i modelli TIN (*Triangulated Irregular Network*) e *mesh*. Questi due tipi hanno funzioni quasi esclusivamente relative alla rappresentazione dell'elevazione sul terreno ed hanno avuto una diffusione più limitata.

2.4.1 Il modello vettoriale

Il modello dei dati vettoriale si basa sulla rappresentazione della realtà attraverso una serie di oggetti grafici di tipo vettoriale. Più che dei veri vettori in senso strettamente matematico, la grafica vettoriale si fonda sull'utilizzo di vertici. Un vertice non è altro che la rappresentazione di una coordinata all'interno del piano cartografico. Così, per la rappresentazione di una retta, si ricorre all'utilizzo di due vertici; uno all'inizio e uno alla fine.

All'interno del modello dei dati vettoriale esistono solo tre tipi possibili di oggetti: punti, linee e poligoni. I punti sono rappresentati da un unico vertice; le linee sono rappresentate da una sequenza di vertici; ed infine i poligoni non sono altro che la rappresentazione di una sequenza di vertici in cui il primo e l'ultimo vengono a coincidere.

[27] Concetto proveniente dall'informatica, utilizzato solitamente per indicare come certi tipi di dati o informazioni, quali i testi, la musica o le immagini, vengano rappresentati all'interno di uno schema puramente numerico come quello delle macchine di calcolo digitale.

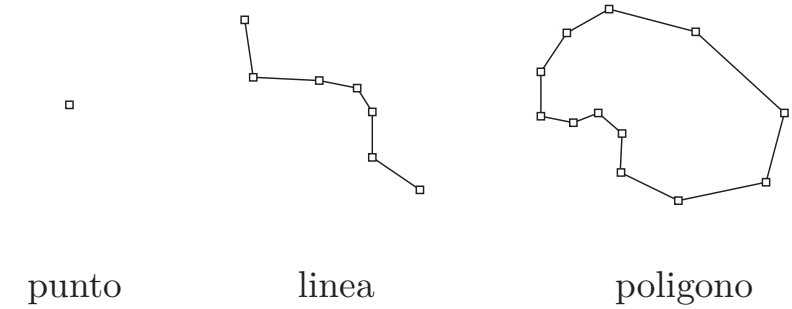


Fig. 2.4. Il modello vettoriale si basa sulla rappresentazione della realtà attraverso tre tipi fondamentali di simboli cartografici: punti, linee e poligoni. A loro volta, queste tre forme di rappresentazione si basano sull'utilizzo di vertici creati grazie al ricorso a coordinate geografiche.

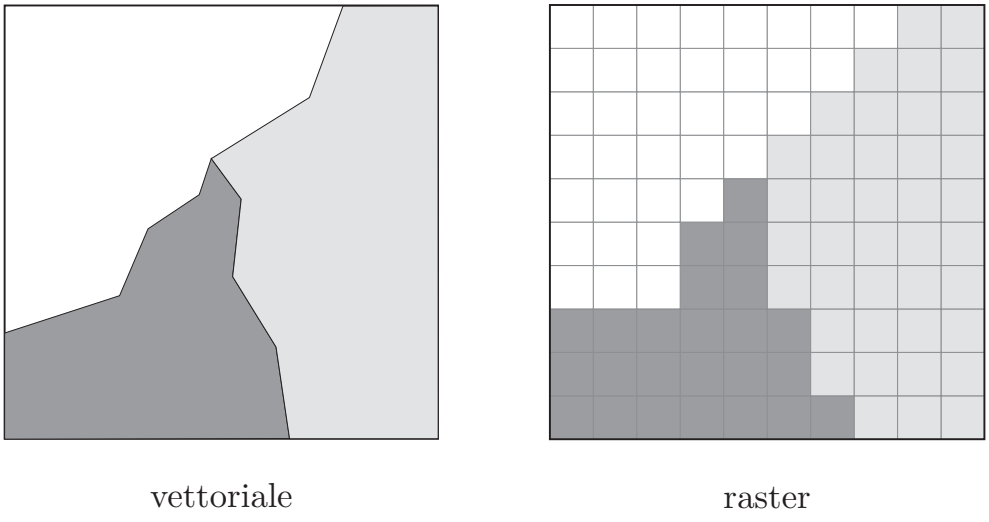


Fig. 2.5. A differenza di quello vettoriale, il modello dei dati *raster* si basa su una griglia o reticolo per rappresentare topologicamente gli attributi spaziali di un determinato territorio. Come si può osservare dalla figura il risultato a livello spaziale è molto diverso. A prima vista il modello *raster* può apparire molto meno preciso. Naturalmente questo è un pregiudizio, dato che basta diminuire la dimensione della cella per ottenere un grado di dettaglio perfino superiore a quello vettoriale.

Si usa il concetto di modello dal momento che la complessità della realtà deve essere per l'appunto sintetizzata con l'aiuto di solo questi tre tipi di elementi geometrici.

Il modello dei dati vettoriale prevede che ad ogni piano cartografico venga associata una tabella dove sono registrati degli "attributi" (dati o informazioni) relativi ad ogni singolo oggetto. In altri termini, ad ogni singolo piano cartografico viene associata una tabella di *database*. Sempre secondo la logica di questo tipo di astrazione della realtà, ad ogni oggetto cartografico corrisponde un unico *record*, secondo uno schema del tipo 1:1.

Il modello dei dati vettoriali ha acquistato nel tempo una grande popolarità, tanto da rappresentare oggi lo *standard* generale per i GIS. Ogni *software* appartenente a questa categoria lavora per prima cosa con piani cartografici di tipo vettoriale. Non solo; la maggior parte degli applicativi a disposizione è costruita attorno alla logica di un modello dei dati vettoriale.

2.4.2 Il modello *raster*

I Sistemi Informativi Territoriali non nascono come applicativi basati sul modello vettoriale. Di fatto, la prima strategia digitale per la rappresentazione cartografica della realtà è stata quella *raster* e i primi *software* GIS si basavano su cartografia di questo tipo. Le immagini o grafica *raster* si basano su una sequenza o matrice bidimensionale di celle (*pixel*) che vanno a ricomporre un'immagine; nel caso del modello *raster* la matrice va a ricomporre un piano cartografico.

Il modello *raster* ha una struttura molto più semplice rispetto a quella vettoriale. Di fatto si tratta di un blocco omogeneo di informazioni. Infatti, la cartografia *raster* non possiede tabelle di *database* associate. L'informazione infatti viene registrata all'interno del *pixel*. La principale differenza tra la grafica *raster* e la cartografia *raster* è che, nel primo caso, all'interno del *pixel* vengono registrati valori cromatici, mentre, nella cartografia *raster*, all'interno del *pixel* vengono registrati dati e informazioni relativi a quella determinata posizione nello spazio. I dati possono essere di qualunque tipo: numerici o alfanumerici e secondo qualsiasi tipo di scala descritta nel paragrafo 4.2.

3
L'analisi locazionale

La definizione di nuovi modelli di localizzazione, così come la loro diffusione e applicazione, è un processo che a più riprese, ha contraddistinto lo sviluppo della geografia nel corso del XX secolo. Questo processo rivoluzionario, che ha visto una rilevante influenza della geografia economica, è stato contraddistinto da due momenti principali. Durante la prima metà del secolo la definizione di nuove teorie e paradigmi geografico-economici; nella seconda, la loro diffusione nelle scuole anglosassoni e del nord Europa, nel processo di esplorazione di nuovi contesti applicativi[1].

Tale corrente si caratterizzò per un marcato approccio formale, basato sulla logica delle leggi economiche. All'interno di questo meccanismo l'uomo appare come un agente sottomesso a regole di mercato e profitto[2]. Le norme economiche introdotte all'interno di modelli spaziali definiscono così dei paradigmi di tipo deduttivo, che, benché espressi come ricostruzioni di nuove forme interpretative della geografia dell'uomo, sono definiti da schemi chiusi causa-effetto. La loro struttura chiusa si estende, ad esempio, nella definizione degli stessi schemi spaziali. Così, se si prende in considerazione il modello di Christaller (vedi paragrafo 3.3) ci si accorge subito che la sua dinamica topologica costituisce a sua volta uno schema chiuso.

Le teoria locazionale[3] delle attività umane può essere considerata come una corrente articolata in più scuole su entrambe le sponde dell'Atlantico che guarda ai processi generali, caratterizzati da meccanismi causa ed effetto all'interno di schemi sociali, politici ed economici. La regolarità non sarebbe altro che un prodotto dei comportamenti umani. In effetti questa corrente, fonda-

[1] Claval 1980, pp. 190-191.

[2] Vagaggini, Dematteis 1976, p. 15: «[...] Questi modelli che ora esamineremo brevemente sono dunque, a stretto rigore, modelli normativi, perché essi non pretendono di illustrare come gli uomini si comportavano nello spazio geografico, ma come essi si devono comportare se vogliono raggiungere certi obiettivi.»

[3] All'interno del presente testo si farà uso del termine locazionale come traduzione dell'inglese *locational*.

ta interamente su modelli, rientra perfettamente nella tradizione *behavourista* che ha caratterizzato tante altre scienze sociali. Questa corrente ambisce a compiere un salto da una descrizione ad una spiegazione del mondo [4].

La figura che ha, forse, contribuito maggiormente alla definizione del perimetro di questa nuova scuola è stato Peter Haggett, in particolare col suo volume *Locational Analysis in Human Geography* nel quale veniva ripercorsa la storia di questa branca e venivano presentati i più influenti modelli teorici oltre a quelli che in modo significativo iniziavano ad affermarsi nel panorama della geografia umana britannica e statunitense[5]. *Locational Analysis* subì una profonda trasformazione con l'edizione integrata con i contributi di Cliff e Frey che portavano così a compimento un'opera molto più estesa ed aggiornata[6]. Indubbiamente il lavoro di Haggett segna un punto fondamentale nell'evoluzione dei modelli locazionali e della loro applicazione nel campo della geografia umana. La versione originale e quella aggiornata sono indubbiamente ambiziose; fatto dimostrato dal tentativo di realizzare un compendio complessivo. Forse proprio per questo motivo, e dato l'elevato numero di tematiche e modelli illustrati, la lettura di *Locational Analysis in Human Geography* richiede in qualche misura una preparazione previa[7]. Il merito dell'opera, oltre che di fornire un compendio cosi vasto, è quello di avere messo a fuoco e ricomposto in modo organico una corrente trasversale composta da più teorie, modelli e metodi geografici che fino ad allora rimanevano sostanzialmente sparsi[8].

In questo senso l'opera di raccolta e ricomposizione di Peter Haggett acquista un significato nuovo. Di fatto, sarà attribuita a lui per la prima volta nella storia della geografia la paternità di un primo paradigma geografico[9]. Secondo Stoddart infatti, saranno Haggett e Chorley ad inserire per primi [in modo confuso: «confusingly»] il concetto di paradigma kuniano. Ed è ovvio che tale principio viene forgiato all'interno della logica locazionale[10].

La tradizione dei modelli economico-geografici nasce – molto tempo prima di Haggett – da un desiderio di spiegare o dare un senso alle dinamiche sociali all'interno dello spazio geografico. Questo movimento cerca di rompere con

[4] Claval 1983, pp. 60-61.

[5] Haggett 1965; traduzione al castigliano in Haggett 1976.

[6] Haggett, Cliff, Frey 1977. Fare attenzione a non confondere le due opere dato che entrambe portano lo stesso nome.

[7] Kolars 1967, p. 276.

[8] Gould 1967, p. 293.

[9] Secondo Kuhn il "paradigma" rappresenta il principio ordinatore di ogni settore della conoscenza: domande, metodi, teorie, prassi ecc. Kuhn 1962. Fare attenzione al fatto che il concetto di "scienza normale" di Kuhn non rappresenta, nell'ambito dell'epistemologia, un concetto universalmente accettato.

[10] Stoddart 1986, pp. 6-11. Chorley, Haggett 1965.

il descrittivismo idiografico tipico delle tradizioni classiche in geografia[11]. In fondo coloro che hanno forgiato i modelli che verranno discussi di seguito, e coloro che ne hanno fatto largo uso considerano che quantomeno una parte della realtà possa essere schematizzata grazie a paradigmi di tipo spaziale. L'analisi locazionale mira così a rispondere a domande naturali della geografia attraverso schemi logico deduttivi; ovvero cerca di non limitarsi a descrivere lo spazio geografico dell'uomo, ma cerca di ricomporlo attraverso l'identificazione delle logiche che ne regolano le forme di occupazione sociale.

Figura emblematica di questa corrente trasversale è Walter Christaller (1893-1969), soprattutto con la sua opera *Die zentralen Orte in Süddeutschland*[12] integrata poi col cosiddetto modello di Lösch e altri ancora[13]. Il luogo di provenienza dei modelli fondamentali è dunque la Germania; Von Thünen, Weber, Christaller e Lösch sono tutti studiosi tedeschi. Sono questi geografi e geografi economici ad avere dato vita a quella corrente che oggi definiamo appunto "teorie normative in geografia", sottodisciplina a cavallo tra la geografia umana e la geografia economica. La nascita e lo sviluppo dell'analisi locazionale hanno previsto, fra le altre cose, un intenso processo di riscoperta di lavori precedenti caratterizzati dall'attenzione da parte degli autori alla distribuzione spaziale dei fenomeni sociali sotto presupposti o schemi comportamentali. Tali lavori erano passati inosservati ed erano rimasti nascosti agli occhi della geografia[14].

Come verrà illustrato nei prossimi paragrafi, i modelli di questi quattro geografi tedeschi vanno ricondotti, attraverso un filo di discendenza o filiazione, ad un unico modello: quello di costo del trasporto. Dal loro studio si potrà capire come sia stato il costo dello spostamento fisico degli uomini e delle merci l'elemento ad avere maggiormente contribuito alla definizione delle relative dinamiche, sia nel caso delle aree di produzione agricola, che in quello delle industrie o delle gerarchie.

[11] Questa corrente geografica rifiuta la visione determinista e possibilista della realtà. Cerca di andare oltre la descrizione del mondo. L'elemento che contraddistingue questa tradizione non è tanto la formalizzazione del discorso geografico, quanto l'identificazione delle trame regolari. Vagaggini, Dematteis 1976, p. 14: «[...] significa porsi interrogativi del genere: i paesaggi e le regioni della geografia umana presentano configurazioni regolari che si ripetono anche in ambienti naturali diversi?».

[12] Christaller 1933.

[13] Lösch (1906-1945) è l'autore di *The Economics of Location* 1959 che costituisce uno dei volumi che maggiormente hanno portato a compimento lo spirito della logica locazionale intesa come processo di comprensione delle strutture normative come teoria delle forme e distribuzione dell'occupazione sociale dello spazio.

[14] Vedi ad esempio il caso della riscoperta di Von Thünen ad opera dello stesso Weber.

Generalmente l'analisi locazionale viene considerata come prodotto, e con esso confusa, del determinismo classico[15]. È vero invece tutto il contrario. I modelli dell'analisi locazionale si basano sì su postulati del tipo causa ed effetto, ma è vero che l'agente che determina i processi o le reazioni non è più l'ambiente, o la natura. L'organizzazione della geografia umana sembra condizionata in primo luogo dalle regole economiche e, in seconda istanza, dalle regole sociali. Entrano così in gioco, in modo definitivo, agenti quali mercato, distanza, beni, servizi, costo di trasporto, collegamenti ecc. Nell'ambito dei modelli locazionali il ruolo determinante lo gioca la distanza, oltre alle condizioni che da essa si generano, quali costi di trasporto, accessibilità, agglomerazione, gerarchia[16]. Prova concreta del fatto che questa branca della geografia poco ha a che fare con il determinismo classico è il fatto che la quasi totalità di tali modelli viene prodotta all'interno di un cosiddetto "spazio ideale". Ovvero uno spazio euclideo ed isotropico: cioè una pianura perfettamente piatta e sconfinata caratterizzata da qualità fisiche e fertilità omogenee. Fattore che determina una identica capacità produttiva e stessi costi di trasporto per unità spaziale su tutta la superficie di questo territorio astratto. La realtà geografica viene così appiattita e privata di tutti i suoi attributi fisici ed ambientali.

Questa tradizione entra in crisi dal momento in cui i modelli in questione si scoprono tuttavia inadeguati alla comprensione reale del mondo. A tale scopo non sono adatti dato che sono troppo semplici e regolari. Si basano su una visione troppo meccanica del mondo. Allo stesso modo, i modelli locazionali non possono essere considerati universali. Essi sono fondati sulla logica del mercato. Dunque difficilmente applicabili in quelle parti del pianeta dove i gruppi sociali hanno sviluppato sistemi economici diversi, dove, ad esempio, vi siano carenza o assenza di una logica del profitto e la presenza invece di una marcata concezione del supporto collettivo dei gruppi sociali o dove si promuova la difesa degli equilibri ambientali.

Lo spazio isotropico rappresenta un modello caratterizzato forse da una sintesi eccessiva della realtà. Nonostante ciò, la modellistica di questa corrente costituisce oggi il percorso propedeutico più efficace per capire come si articolino i processi umani all'interno dello spazio, dato che questi introducono in modo chiaro ed elementare concetti fondamentali per la geografia quali "posi-

[15] Quello di Ratzel e Churchill Semple; Churchill Semple 1908.
[16] Tinacci Mossello 1990, pp. 91-92; Vallega 2004, pp. 23-24.

zione", "regione", o "territorio"[17]. Il lettore sappia tenere conto, in ogni caso, della portata parziale dei concetti che verranno introdotti di seguito. A prescindere da tutto ciò, si tenga anche presente che questi modelli seguono un filo di sviluppo. Pur essendo stati prodotti in modo da rispondere a domande diverse, ognuno di loro ha contribuito nel fornire elementi che sono serviti a migliorare quello successivo.

Pur essendo stati fondati inoltre come ambiziose teorie generali (se non universali) relative alla geografia della specie umana, e ripercorrendo parte della storia e della logica dell'applicazione delle teorie locazionali, si può osservare che questi modelli hanno un ruolo importante e un'applicabilità nel campo delle ricerche regionali[18].

In essi lo spazio viene appiattito creando così una rappresentazione semplificata della realtà fisica[19]. In fondo si parla di "modelli economico-spaziali", più che geografici. Ed è qui il punto che si desidera mettere in evidenza. Lo spazio – quello del geografo – costituisce, effettivamente, un concetto fin troppo astratto e complesso, caratterizzato da multidimensionalità, deformazioni e percezioni[20]. In estrema sintesi, lo spazio appartiene a quella categoria di cose e concetti che possono essere appresi o imparati, ma non insegnati. I modelli geografici sono appunto una scorciatoia per agevolare il processo di apprendimento di un oggetto di studio così complesso. Essi, oltre ad insegnarci una serie di principi fondamentali relativi alle leggi di mercato, e all'economia delle società occidentali, presentano la dinamica spaziale in modo chiaro e diretto.

Per quanto possa apparire superfluo farlo, è importante ricordare come le teorie locazionali rappresentino quanto meno un insieme di idee comuni ad una intera comunità di ricercatori. All'interno della dialettica sullo spazio i concetti di "fascia rurale", "area di pertinenza" o "area di mercato" possono essere sfruttati in modo preciso per rappresentare elementi specifici di una realtà che altrimenti potrebbe apparire innecessariamente caotica e soggettiva.

[17] Questo requisito propedeutico è da sempre stata una necessità della geografia. Vedi ad esempio Hartshorne 1939, p. 36 quando riferendosi al periodo "pre-classico" della geografia afferma: «Physical geography is the essential propaedeutic for an understanding of our perceptions of the world, whether those received directly by travel or indirectly by reading [...]. In brief, "it serves as a suitable arrangement for our perceptions, contributes to our intellectual pleasure, and provides rich material for social disclosure." [Nota a Rink].»

[18] Sulla contrapposizione tra modelli normativi classici e il «rinnovamento degli studi regionali» consultare Claval 1996, pp. 21-23.

[19] Sul tema della natura semplicistica dei modelli e di come questi possano essere riorganizzati in schemi più complessi consultare Boots 1980; Lloyd, Dicken 1986, pp. 33-34.

[20] Aase 1994; Couclelis, Gale 1986: «Thus we can say that at least under one set of formal criteria, cognitive space is qualitatively different from perceptual space.»

Ma l'aspetto più importante relativo al quadro delle teorie normative riguarda la loro applicabilità nel quadro degli studi regionali. Fattore che va esteso dunque al di là dei confini della geografia nel campo delle ricerche delle diverse discipline umane interessate alla comprensione dello spazio. Ci si limita in questa sede a presentare quei modelli originali che, in un modo o nell'altro, possono contribuire, attraverso l'elaborazione di nuovi paradigmi o processi di comparazione, al miglioramento della nostre conoscenze sulle forme della geografia umana e sulla loro storia.

3.1 I poligoni di Thiessen o di Voronoy

I "poligoni di Thiessen" non possono essere considerati come un modello geografico in senso stretto; per molti infatti questi appartengono anche alla categoria dell'analisi spaziale e non di quella locazionale. Nonostante ciò, sarà il primo modello qui presentato. Nell'ambito della geografia umana, questo modello, comunemente conosciuto come appunto i poligoni di Thiessen[21], in sostanza risponde alla domanda "dove si collocano i confini dell'area di pertinenza ideale di un determinato insediamento?" Per area di pertinenza si intende il territorio meglio controllabile da un insediamento o, in altre parole, il territorio che gli spetta sulla base della suddivisione ideale dello spazio. La sua importanza naturalmente va nella direzione della definizione di processi di controllo politico del territorio.

I poligoni di Thiessen hanno un'origine molto più complessa di quanto il nome unico generalmente usato possa suggerire; infatti sono almeno tre i modi di chiamarli. In mero ordine cronologico, il primo ad avere definito il paradigma topologico alla base di questo modello geografico fu Dirichlet[22], matematico tedesco al quale si deve, fra le altre cose, anche la definizione formale di "funzione matematica".

In sostanza, il ragionamento di Dirichlet era quello di definire a livello geometrico il frazionamento di un piano a partire da una serie di punti seguendo una serie di regole specifiche[23]:

- lo spazio può essere suddiviso solo in poligoni convessi; questo comporta che tra questi sia definita una trama ordinata che permetta ad ogni porzione di mantenere la forma convessa;

[21] Sul valore e il ruolo dei poligoni di Thiessen come modello nel campo della geografia umana consultare Garrison 1959, pp. 471-482.

[22] In ambito matematico e topologico, i poligoni di Thiessen sono appunto infatti noti anche come *Dirichlet tessellation* o "trame di Dirichlet", formulate nel 1850 a livello formale per la prima volta da Johann Peter Gustav Lejeune Dirichlet (1805-1859).

[23] Sulla costruzione dei poligoni di Thiessen si rimanda a Kopec 1963. Consultare anche Boots 1987.

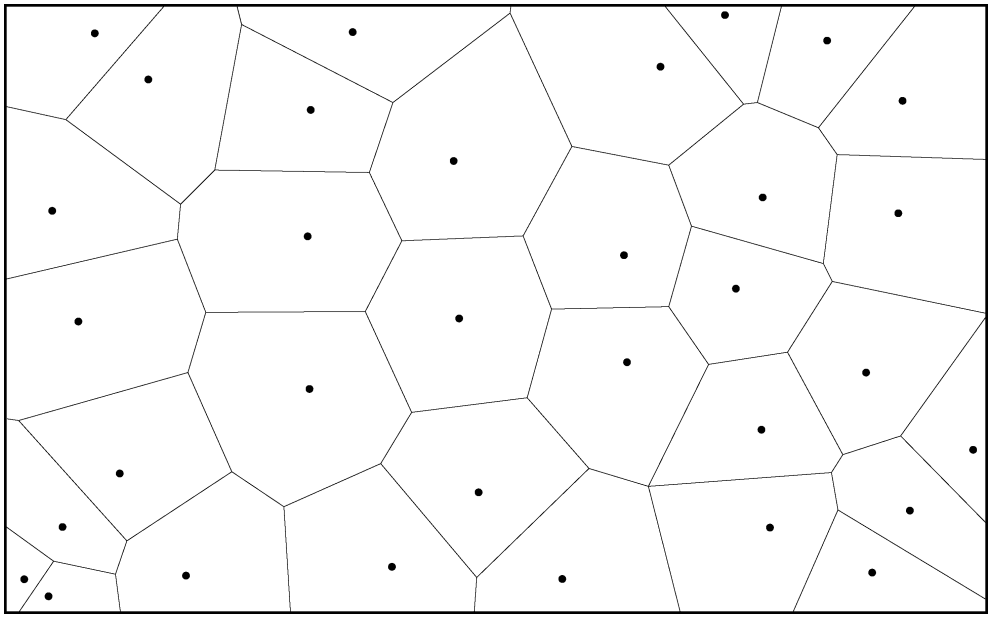


Fig. 3.1. I poligoni di Thiessen rappresentano, per una serie di "punti generatori", le aree di pertinenza (ovvero il territorio meglio controllabile) di ciascuno determinate all'interno di uno spazio euclideo ed isotropico.

- i poligono devono essere definiti a partire da una serie di "punti generatori". Ognuno dei poligoni può contenere uno solo dei punti generatori;
- il poligono generato dovrebbe essere tale che ogni posizione al suo interno dovrebbe essere più prossima al proprio punto generatore che a qualunque altro.

Il risultato concreto di tale procedura astratta è rappresentato nella figura 3.1. La proprietà principale di questi poligoni è dunque che ogni loro parte o porzione è più vicina al proprio punto generatore che a qualunque altro. Detto in altri termini, ogni poligono delimita lo spazio più prossimo al proprio punto generatore.

Tale idea venne ripresa non tanto tempo dopo da Georgy Voronoy, matematico russo che estese la notazione e la definizione formale delle aree. Da quel momento questi poligoni vengono conosciuti anche come le "regioni di Voronoy" (o Voronoi). Fin qui niente che potesse essere minimamente ricondotto alla geografia. In effetti sia Dirichlet che Voronoy si limitavano all'ambito disciplinare della topologia e della matematica.

Nel 1911 un meteorologo del *Weather Bureau Service* degli Stati Uniti d'America pubblicò un breve articolo nel quale esponeva una nuova metodologia

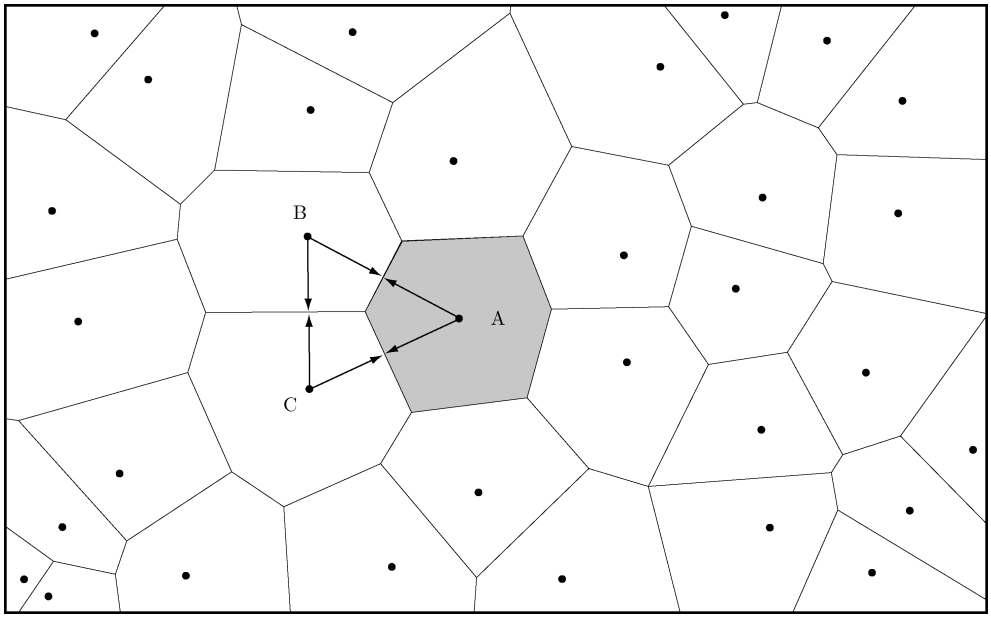


Fig. 3.2. Nella figura si può osservare il poligono (di Thiessen) grigio generato a partire dal punto A. La sua caratteristica più importante è che ogni porzione di esso è più vicina ad A che a qualsiasi altro punto generatore.

per la previsione delle precipitazioni[24]. Il suo nome era Alfred H. Thiessen. Benché il suo saggio non avesse riscosso grande successo nell'ambito della meteorologia, si trattava della prima volta che si faceva riferimento all'uso delle trame di Dirichlet o delle regioni di Voronoy nella realtà geografica. Da allora gli usi di tali suddivisioni sono stati i più svariati e hanno comportato un ricorso costante ad essi in diversi settori della geografia, in particolare, naturalmente, nell'ambito della geografia umana. Perciò, anche se Thiessen non è stato effettivamente il primo ad avere definito la struttura e la natura di questo schema geometrico, si fa proprio riferimento a lui in relazione all'uso e all'applicazione di essi.

L'importanza di questo modello nell'ambito della geografia umana va intesa nel senso di definizione dello spazio primario d'interazione sociale. Naturalmente la sua applicazione anche in campo geografico è alquanto limitata e resta tutto sommato nell'astratto, ma diviene fondamentale anche per la costruzione e comprensione di altri modelli geografici che verranno trattati più avanti. Per molti versi si tratta del primo passo per addentrarsi nella logica

[24] Thiessen 1911.

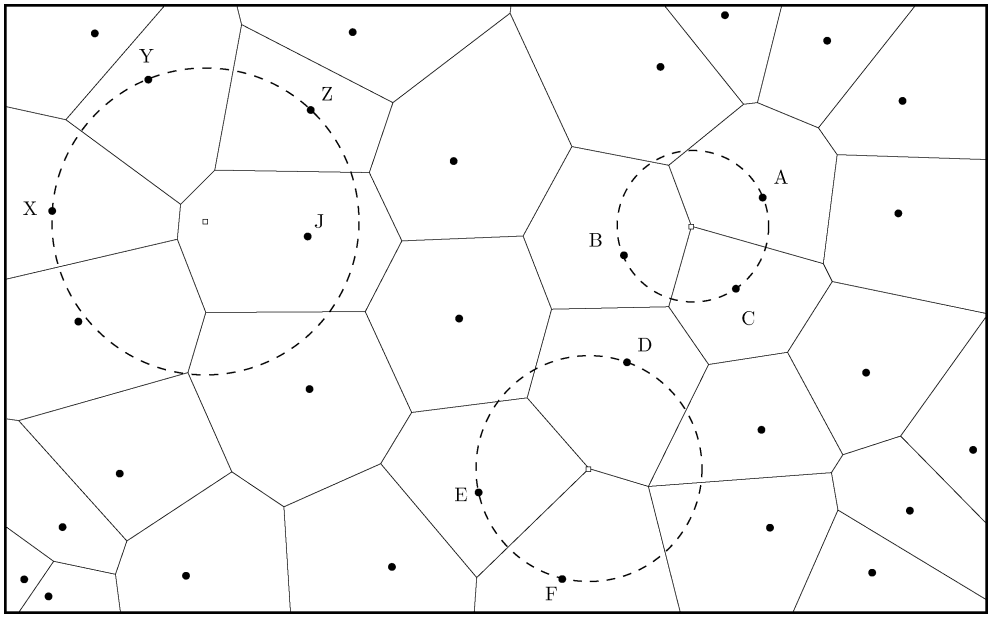


Fig. 3.3. La procedura più ortodossa per la definizione dei poligoni di Thiessen richiede l'identificazione dei vicini geografici di un determinato punto.

dell'analisi locazionale, e perciò va considerato essenziale per la comprensione delle strutture spaziali intese come paradigma. Lo studente nell'osservare questo modello per la prima volta può cogliere gli elementi primari del peso coercitivo dello spazio sulle forme d'occupazione sociale dello stesso.

Come la quasi totalità dei modelli locazionali, i poligoni di Thiessen vengono generati all'interno di uno spazio euclideo ed isotropico; da ciò la valenza astratta di cui si diceva prima. Nel campo della geografia umana si è soliti definire tali poligoni come "territori di pertinenza" o "aree di pertinenza". Immaginando un territorio delimitato da confini all'interno del quale sono presenti insediamenti simili, ovvero di uguale rango gerarchico[25], in termini di dimensione e di funzioni, il proprio poligono rappresenterà, per ogni singolo centro, il territorio immediatamente e più facilmente controllabile; perché comprendente tutte le frazioni di terra più vicine a quel centro. Ad esempio, immaginando due contadini residenti in due insediamenti vicini, possiamo considerare le terre delimitate dal poligono come quelle più facilmente raggiungibili da ciascuno di loro. Se decidessero di partire alla stessa ora del mattino,

[25] Il concetto di rango gerarchico verrà trattato nel paragrafo relativo al modello dei "luoghi centrali" di W. Christaller.

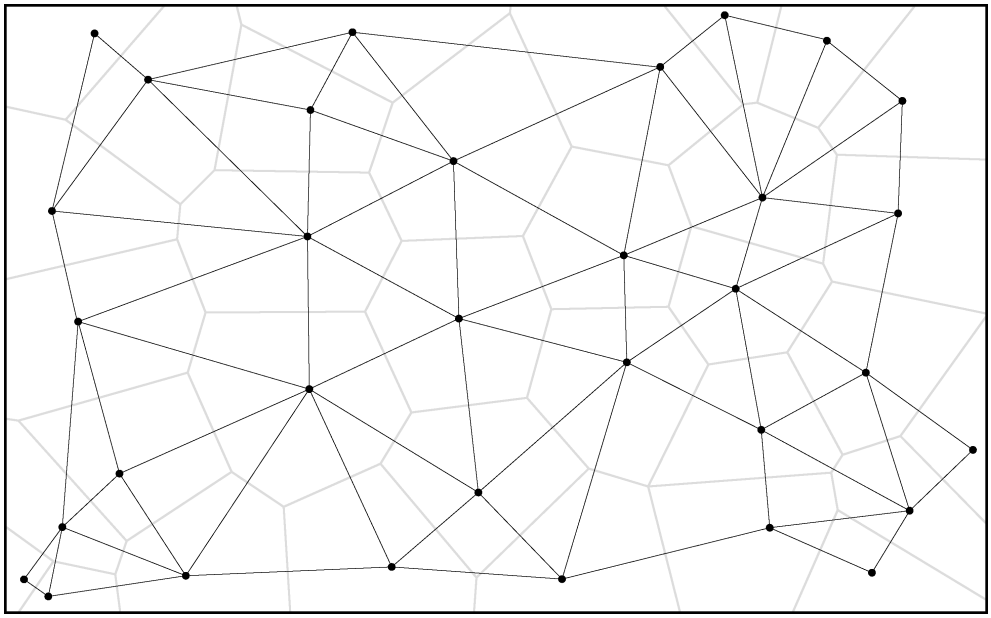


Fig. 3.4. Rappresentazione delle trame di Delaunay.

ognuno nella direzione dell'altro, finirebbero per incontrarsi a metà strada. Infatti una delle caratteristiche principali dei poligoni di Thiessen è che il confine di due insediamenti vicini si muove su un asse perpendicolare alla retta che li separa. Non esiste porzione di terra all'interno del poligono di un centro che i residenti in altri centri possano raggiungere prima.

Il poligono, in senso generale, viene ad assumere così il significato di confine territoriale. Nella figura 3.5, è rappresentato il confronto tra i confini dei comuni della provincia di Siena e i poligoni di Thiessen generati a partire dai capoluoghi. Come si può osservare, i volumi e l'ingombro cartografico di questi rispecchiano in qualche misura quelli delle aree amministrative. Queste proprietà sono state sfruttate da diversi studiosi delle civiltà più antiche nel processo di ricostruzione dei confini in quei contesti per i quali, per diversi motivi, si è persa traccia della reale composizione dei quadri di potere[26].

Vi è inoltre un aspetto che non viene solitamente preso in considerazione nella trattazione relativa ai poligoni di Thiessen: il concetto di "vicinanza geografica". In effetti rappresenta un argomento che solitamente viene preso in considerazione nella trattazione relativa alla costruzione dei poligoni; ovvero

[26] Recenti applicazioni del modello di "pertinenza" si possono trovare in D'Ercole 2000; Savage, Falconer 2003 p. 38.

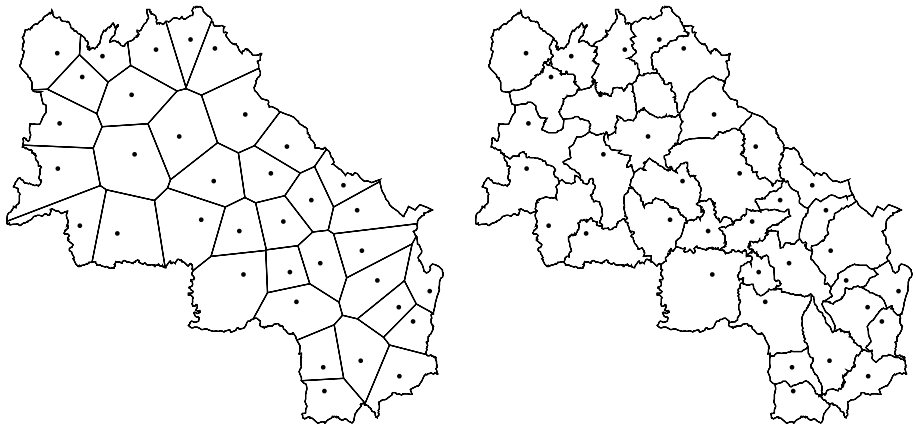


Fig. 3.5. Confini amministrativi dei comuni della provincia di Siena e rappresentazione cartografica dei poligoni di Thiessen generati a partire dai capoluoghi di comune.

relegata nella letteratura tecnico-informatica[27]. A prescindere da tutto ciò, occorre ritornare sul tema della vicinanza geografica dato che esso non solo si presta a far capire meglio l'articolazione e la strutturazione dello spazio umano, ma consente di capire anche altri aspetti che verranno trattati di seguito.

In sostanza, la vicinanza geografica indica quando due insediamenti possono essere considerati "vicini spaziali". È stato già chiarito che un poligono di Thiessen possiede confini esclusivamente con quelli dei vicini geografici (ovvero non con quelli che a una lettura solo visiva della carta possano apparire vicini); dunque appare lecito domandarsi cosa sia effettivamente la vicinanza geografica. Il concetto si riallaccia ancora una volta a principi geometrici e topologici. Questi però non ci permettono di identificare quando due punti (o insediamenti, nel nostro caso) sono vicini; per affermarlo occorre che i punti siano tre. Esiste di fatto una regola geometrica che afferma come dati tre punti, esiste uno ed un unico cerchio per i quali passa la sua circonferenza. Nella figura 3.3 si può osservare (linea tratteggiata) il cerchio che passa sui punti A, B e C. Se, come in questo caso, nessun altro punto cade all'interno del cerchio tracciato, dunque, i tre punti possono essere considerati vicini. Viceversa, se tracciando un cerchio la cui circonferenza passa per tre punti (come

[27] Per una descrizione delle problematiche tecniche e per una bibliografia completa consultare Tsai 1993.

per esempio nella figura 3.3, X, Y e Z), uno o più punti della distribuzione analizzata cadranno al suo interno (come ad esempio J), allora quei tre punti non potranno essere più trattati come vicini.

Questo non significa che una coppia di punti non possa dimostrare la sua vicinanza con un altro vicino (ad esempio i punti X e Y lo sono rispetto al cerchio tracciabile su XYJ). Significa solo che quei tre punti (XYZ), intesi come insieme, non dimostrano di essere vicini. Nella figura 3.3 i dati illustrati mostrano come il cerchio tratteggiato confermi la vicinanza geometrica dei punti A, B e C. Non solo; il centro del cerchio possiede una proprietà particolare. Infatti, come si può osservare, il suo centro va a collocarsi in corrispondenza del vertice comune dei tre poligoni di Thiessen dei vicini. Il principio di vicinanza spaziale è fondamentale dato che permette di discernere tra punti che effettivamente sono o non sono vicini. In effetti, in molti casi, vi sono insediamenti che possono apparire troppo lontani per potere essere considerati vicini, eppure, nonostante la lontananza, risultano esserlo. Si osservi il caso dei punti D ed F: apparentemente troppo lontani, eppure vicini.

Il concetto di vicinanza geografica permette di introdurre un ulteriore aspetto legato in qualche misura ai poligoni di Thiessen: la rete o triangolazione di Delaunay[28]. In effetti, dato su un territorio un determinato numero di punti, sarà lecito domandarsi quale sia il miglior modo di collegarli attraverso un insieme di rette. Colui che segue un percorso arbitrario presto si renderà conto che molte sono le combinazioni possibili. Eppure dovrebbe esistere un modo coerente, e, per così dire, naturale di collegare i punti. La rete di Delauney è composta esclusivamente dai collegamenti tra punti che rispettano il principio di vicinanza spaziale. Questa rete possiede diverse proprietà, tra le quali (e forse la più importante per il geografo) quella della massima efficienza.

3.1.1 «Lo stato isolato» di Von Thünen

Johann Heinrich Von Thünen (1783-1850) è stato la figura che in modo più rilevante ha influenzato l'evoluzione delle teorie locazionali nella geografia. Il suo modello infatti, noto come "teoria dello stato isolato", avrebbe definitivamente caratterizzato sia il linguaggio che la logica con la quale nel XX secolo i geografi avrebbero costruito i propri modelli geografici.

Von Thünen in giovane età acquistò la tenuta di *Tellow* (nella regione di Meclemburgo) che gestì con un certo successo per più di quarant'anni, fino alla sua morte[29]. Le sue doti di contabile e organizzatore lo portarono ad acquisire

[28] Per una descrizione formale della trama di Delauney, fare riferimento a Miles 1972.

[29] Per quanto concerne la "teoria dello stato isolato" si consiglia come saggio introduttivo Grotewold 1959.

nel tempo – in modo diretto e grazie all'esperienza sul campo – una estesa serie di dati e resoconti relativi a costi di produzione, resa e profitti riferiti alla posizione o localizzazione dei diversi terreni agricoli e aree produttive. In sintesi, il suo lavoro è stato quello di correlare in un'unica teoria spaziale le informazioni raccolte lungo il suo percorso di amministratore[30]. Sintesi del suo lavoro è stata appunto la teoria, o modello, nota come "stato isolato", dal titolo della sua opera, pubblicata in tre parti, *Der isolierte Staat*[31]. Si tratta di un modello ideale e astratto che ha come scopo primario definire o dare una spiegazione alla distribuzione spaziale dei diversi tipi di produzione agricola[32].

Secondo Von Thünen, osservando idealmente gli spazi produttivi attorno a un'area urbana, si potrebbe rilevare una sorta di specializzazione caratterizzata da una serie di anelli concentrici attorno al nucleo. Nelle zone più vicine al centro dello stato isolato si osserverebbe un'area dedicata alla produzione di ortaggi e latticini, dopo la quale ci sarebbe un anello di foreste per la produzione di legname. A queste prime due aree seguirebbero più anelli, relativi, a seconda del tipo di rotazione, a diversi tipi di coltivazioni, seguiti da una fascia dedicata all'allevamento di bestiame da macellazione e infine da uno stretto anello di terre incolte.

La sequenza delle fasce, a prima vista arbitraria e casuale, sarebbe legata a principi fondamentali alla base della loro successione e organizzazione spaziale. Ad esempio, gli ortaggi e i latticini verrebbero prodotti nelle aree immediatamente vicine alla città dato che, per questo tipo di prodotto, l'aspetto predominante è la conservazione e dunque la breve durata del prodotto. Ortaggi e latte (così come alcuni dei suoi derivati quali burro e altri latticini freschi) presentavano, prima dell'introduzione dei sistemi di refrigerazione e dei sistemi di pastorizzazione, tempi di conservazione minimi. Fattore che comportava la necessità che fossero consumati quasi subito dopo la loro raccolta o preparazione. L'unica soluzione pratica era dunque la produzione di questi beni nelle aree immediatamente contigue al mercato cittadino.

Già da questo piccolo esempio appare chiaro che l'elemento più rilevante nella logica della successione degli anelli di Von Thünen è la distanza dal centro, o sarebbe meglio dire, dal mercato cittadino; perché è là, secondo la logica di questo modello, che gli scambi determinano la conformazione degli anelli[33]. Sarebbe dunque l'accessibilità al mercato o, meglio ancora, i costi di

[30] Tinacci Mossello 1990, pp. 253-254.

[31] Von Thünen 1826.

[32] Un'altra definizione efficace e chiarificatrice dell'essenza di questo paradigma è «modello elementare della rendita dei suoli»: vedi Poli 2001 p. 55.

[33] Claval 1983, p. 204.

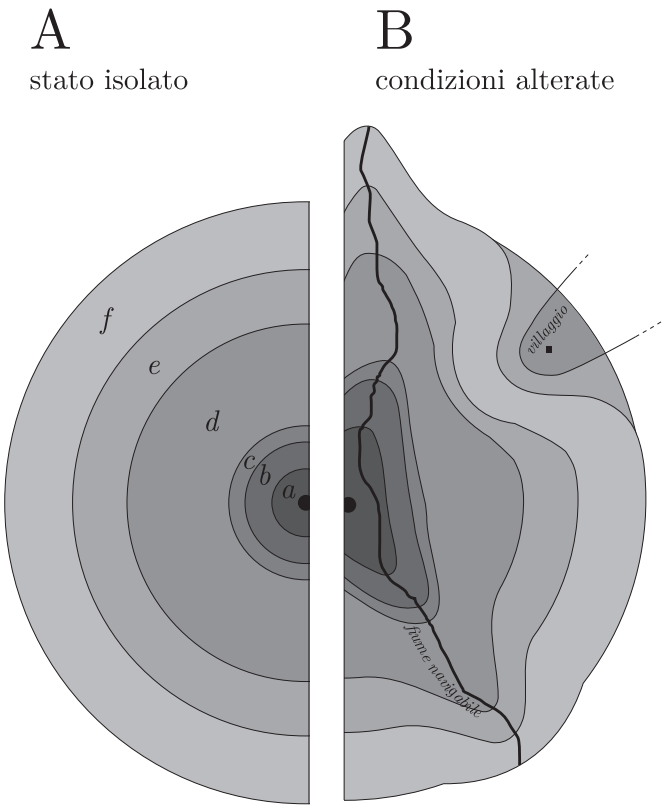


Fig. 3.6. Secondo Von Thünen, in uno spazio ideale caratterizzato dalla presenza di un unico centro abitato dove effettuare gli scambi, la distribuzione dei diversi tipi di specializzazione agricola avrebbe assunto una conformazione ad anelli concentrici, come illustrato nello schema A. La tipologia delle produzione nei diversi anelli è: *a* latticini e ortaggi; *b* legname; *c* agricoltura intensiva *d* agricoltura estensiva *e* rotazione triennale *f* allevamento bestiame da macellazione. Lo schema B illustra invece come la perfezione e la simmetria del modello venga rotta con semplici modifiche quali la presenza di nuovi centri di scambio o vie di comunicazione.

trasporto a stabilire, in base alla localizzazione, il tipo di produzione ideale sulla base di un principio di profitto[34].

Seguirebbe dunque il legname. In questo caso il problema non è la conservazione ma i costi di trasporto del bene. Dato che si tratta di un prodotto molto pesante, far crescere foreste in aree remote comporterebbe costi di trasporto difficilmente sostenibili da un'economia di mercato. Seguirebbero diversi anelli di produzione relativa a coltivazione estensiva. Questi prodotti hanno dei costi di trasporto inferiori a parità di volume in relazione al legname. Segue infine l'allevamento di bestiame e la pastorizia. Questo tipo di attività viene relegata alle aree più remote proprio per il fatto che i capi allevati si muovono autonomamente. Per questo motivo il costo di trasporto è inferiore rispetto a quello di altri tipi di produzione.

La teoria si basa su una serie di presupposti fondamentali che servono da premessa per le successive deduzioni:

- il concetto alla base del modello di Von Thünen è quello, come indica il suo stesso nome, di uno stato completamente isolato. Questo stato non dispone di contatti di nessun tipo con altre forme statali;
- questo stato ideale possiede un unico centro urbano, posto esattamente al suo centro;
- nello stato ideale non esistono altre forme di stanziamento. Questo significa che nel modello di Von Thünen, l'unico luogo di residenza è la città posta al centro dello stato;
- il modello non prevede la presenza di strade o altre vie di comunicazione;
- un altro elemento chiave del modello è che la geografia fisica dello stato isolato equivale ad uno spazio euclideo ed isotropico. Cioè non esistono barriere fisiche che possono impedire il trasporto o il movimento al suo interno. I costi e i tempi di spostamento sono omogenei, ovvero, a pari distanza, essi si equivalgono;
- sempre nel modello di Von Thünen non esistono differenze nella qualità dei suoli in termini di qualità o resa.

L'ordine degli anelli concentrici avrebbe avuto delle caratteristiche specifiche fondamentalmente determinate da cinque fattori principali:

- E costo di produzione di un prodotto agricolo specifico.
- f costo di trasporto di quel determinato prodotto agricolo per tonellata.
- p costo di mercato dei prodotti agricoli.
- k distanza del luogo di produzione dal mercato.
- Y ricavo in tonnellate per ettaro del prodotto agricolo in questione.

[34] Haggett 1967, p. 13.

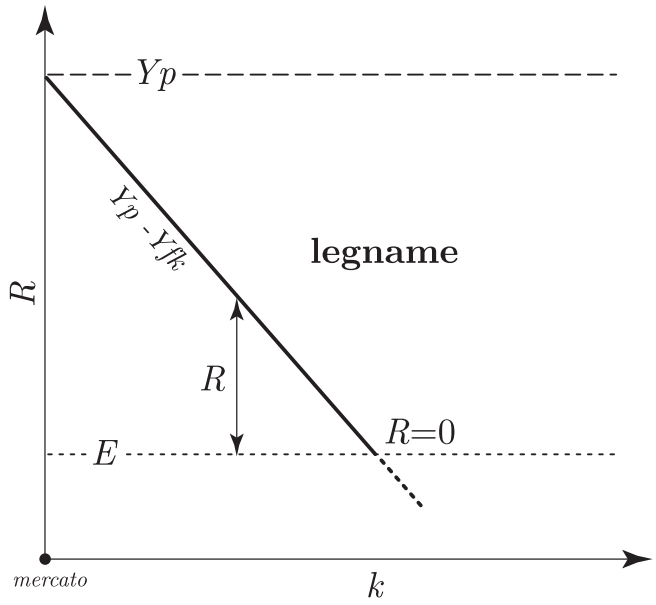


Fig. 3.7. Se si prende il caso della produzione del legname, si può osservare come i profitti dopo la vendita si collochino all'interno di un intervallo compreso tra i costi di produzione (E) e i prezzi di vendita (Yp). All'interno di questo intervallo l'unico elemento che varia sono i costi di trasporto dal luogo di produzione al mercato (Yfk). Questi determinano la graduale riduzione dei profitti (R) in relazione al progressivo allontanamento dal mercato.

Secondo Von Thünen, l'elemento chiave per comprendere il motivo per il quale le aree produttive assumono all'interno dello stato isolato la conformazione circolare, sono i redditi R possibili delle singole unità produttive. Tale resa va calcolata come:

$$R = Yp - E - Yfk \tag{3.1}$$

o, più semplicemente, i redditi (R) equivalgono al profitto lordo da vendita nella piazza del mercato (Yp), meno i costi di produzione del bene (E), meno i costi di trasporto del bene dal luogo di produzione al mercato (Yfk).

Prendiamo ad esempio la produzione di legname, ovvero la specializzazione produttiva del secondo anello intorno alla città. La rappresentazione grafica dell'applicazione dell'equazione 3.1 può essere osservata nella figura 3.7. Come si può leggere nel grafico, il costo di produzione (E) è rappresentato come una costante dalla linea punteggiata. Qualora il ricavato finale del legname R risulti inferiore al costo di produzione E, quel tipo di produzione diviene non

sostenibile nel senso che non solo non permetterebbe di ottenere dei guadagni, ma comporterebbe delle perdite. In altre parole, esiste una distanza oltre la quale non conviene produrre il legname. Nello stesso modo il grafico mostra come il limite massimo dei possibili guadagni R si ottiene qualora il bene venga prodotto a ridosso o nell'immediata vicinanza della piazza del mercato. Man mano ci si allontana da questa, il costo di trasporto Yfk cresce, determinando un abbassamento di R.

Il grafico mette in evidenza come il fattore che maggiormente determina la variazione dei profitti derivanti dalla produzione di un certo prodotto agricolo in base alla posizione, sia la distanza k dalla piazza del mercato. La curva dei guadagni va rappresentata come una retta che va dal punto di massimo a quello di minimo profitto. Oltre questo intervallo, rappresentato dalle due rette parallele, tratteggiate, la produzione di questo bene è impossibile.

Sulla base di queste considerazione, è possibile giungere all'esposizione definitiva del modello. Ogni tipo di produzione possiede una retta R che rappresenta la variazione dei guadagni sulla base della distanza dal mercato. Per ogni singola retta cambiano naturalmente i costi di produzione E e quelli di trasporto f. Questo fatto determina che la retta R per ogni tipo di produzione presenterà un diverso tipo d'inclinazione, come illustrato nella figura 3.8.

Sempre in questa immagine si può osservare anche che tutte le rette vanno a sovrapporsi o a incrociarsi reciprocamente. La domanda alla quale il modello di Von Thünen dà risposta è proprio l'identificazione dei confini tra un anello di produzione e l'altro. Osservando la figura 3.8 si può notare come la retta della produzione di ortaggi e latticini incroci quella del legname. C'è una distanza k oltre la quale risulta più conveniente, in termini di profitti netti, la produzione di legname che non quella di ortaggi. Il punto esatto, quello dove le due rette si incrociano, viene a configurarsi come il confine tra ortaggi e legname.

Il modello dello stato isolato è stato applicato in più contesti da geografi di tutto il mondo per verificare la sua validità. Uno dei casi più noti è quello di Horvath che cercò di testare l'attendibilità della teoria di Von Thünen nel caso di una città africana (Addis Abeba) dimostrando un certo livello di conformità e «parallellismo» con il modello originale[35]. Un altro esempio molto importante è quello proposto da Griffin, la cui applicazione del modello, pur mettendo a fuoco il ruolo centrale di Montevideo, estende i principi locazionali del modello su base regionale. Anche in questo caso emergono delle coincidenze significative, pur distinguendo in modo critico come la differenziazione dei suoli, fertilità, morfologia o altri fattori come strade e collegamenti finiscano per deformare la struttura originale[36].

[35] Horvath 1969, p. 308.
[36] Griffin 1973, pp. 508-510.

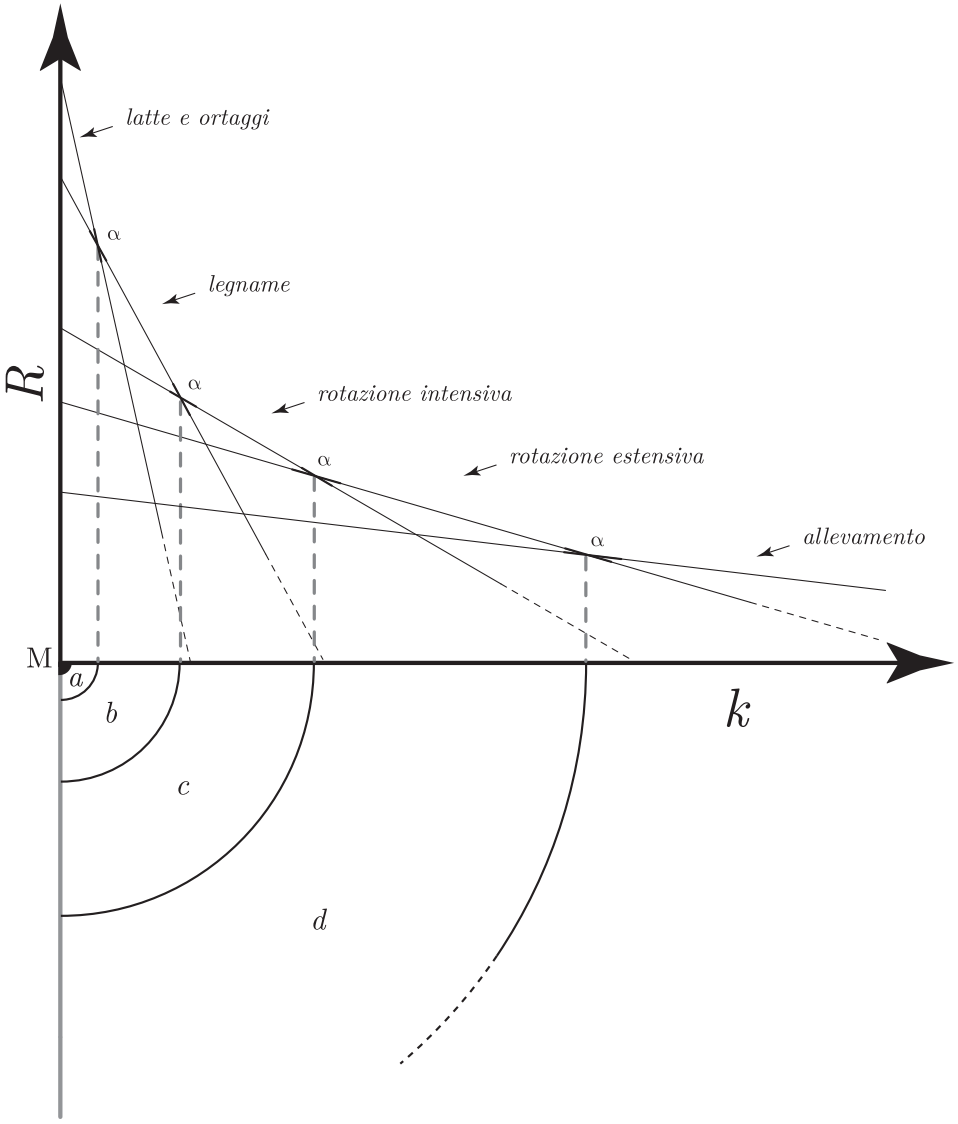


Fig. 3.8. Sovrapponendo le rette di profitto si può osservare come queste, partendo dall'asse verticale, presentino un angolo d'inclinazione determinato dai costi di trasporto. Partendo da sinistra, i profitti per la produzione di ortaggi sono maggiori di quelli ricavati dal legname. C'è un punto però (segnato da α) in cui scendono sotto quello del legname; e uno dove scendono quelli della rotazione intensiva. Le sovrapposizioni α segnano i punti in cui idealmente sull'asse k si collocano i confini tra gli anelli di produzione.

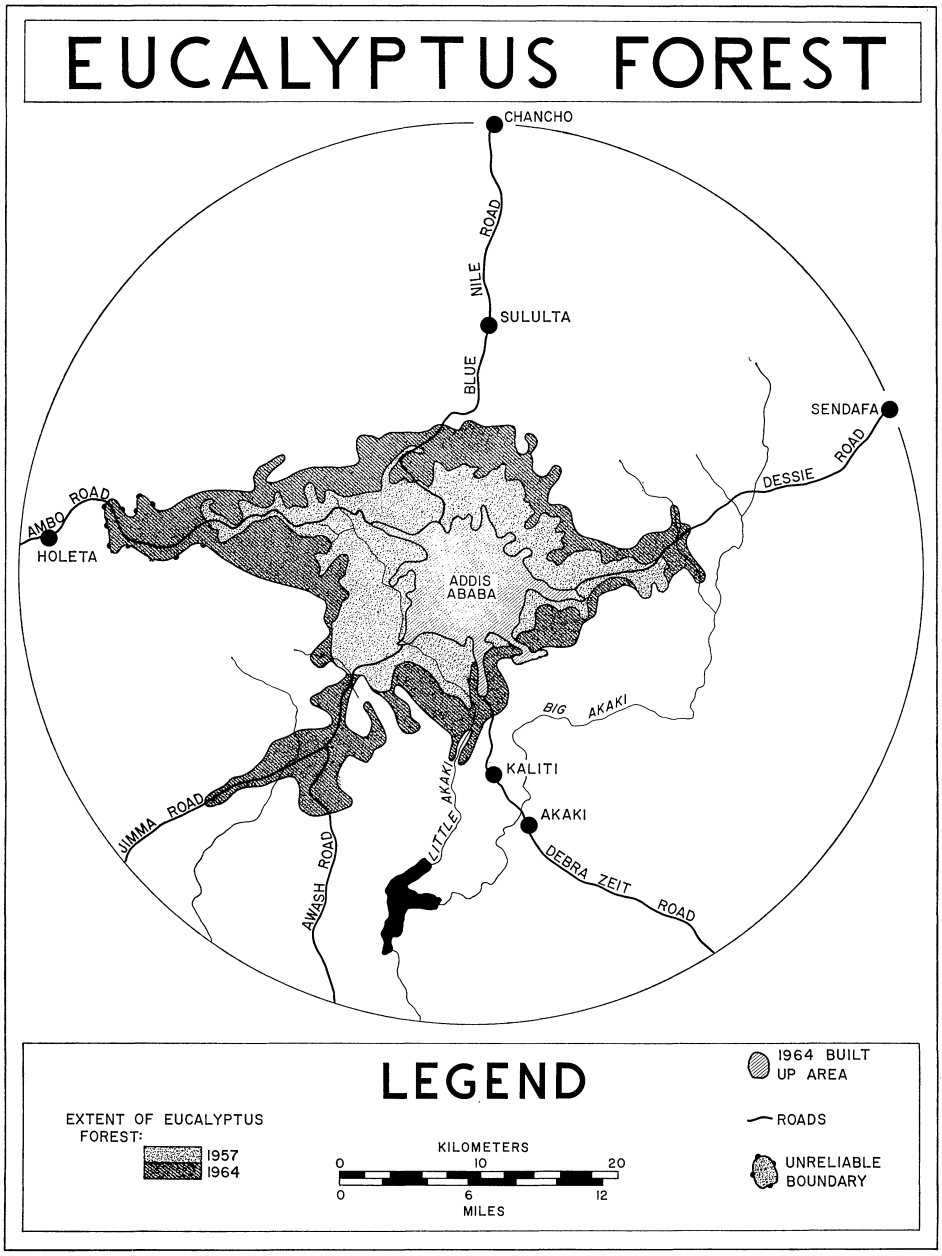


Fig. 3.9. Da Horvath 1969. Anelli di produzione di eucalipto attorno al centro urbano di Addis Abeba.

Concludendo, il modello di J.H. Von Thünen, nella sua grande semplicità e astrazione, ha dimostrato più volte la sua validità, se non altro nell'identificazione dei *patterns* locazionali generati per effetto dei grandi centri urbani e le dinamiche di scambio e consumi al loro interno.

3.2 La teoria della localizzazione delle industrie

Nel 1909 Alfred Weber (1868-1958, economista e sociologo, fratello di Max Weber) pubblicava la sua opera *Über den Standort der Industrie*[37] nella quale esponeva una teoria sui principi fondamentali e i fattori che giocavano un ruolo determinante nella scelta delle aree di produzione industriale[38].

Ad una prima lettura del testo di questo autore, si intuisce subito il ruolo di rilievo che la logica di Von Thünen ha avuto nell'ispirazione del pensiero di Weber[39]. Di fatto, come verrà illustrato di seguito, nel caso della produzione industriale come in quello della specializzazione produttiva delle aree agricole, il ruolo determinante lo giocano ancora una volta i costi di trasporto[40]. Naturalmente la sua teoria veniva elaborata in un periodo di grande rivoluzione e trasformazione, nel quale le nuove necessità dell'industria e la società richiedevano un'attenta e scrupolosa analisi del problema della localizzazione delle aree di produzione. D'altra parte il modello di Weber andava oltre la produzione industriale, e cercava di definire le posizioni ideali anche di altri tipi di attività.

Il modello di Weber nasce dalla definizione di tre diverse tipologie di luogo.

- Le località R, in cui vengono estratte le materie prime.
- Le località P, dove avviene il processo di trasformazione delle materie prime.
- Le località M, o più semplicemente i mercati dove i prodotti dell'attività industriale vengono scambiati come beni.

Le tre tipologie di luoghi rispondono dunque a necessità e principi locazionali specifici. Le aree di estrazione possono essere solo lì dove le risorse si trovano. Nello stesso modo, i mercati si troveranno al centro dei nuclei e stanziamenti con una dimensione consistente. La domanda a cui la teoria della localizzazione industriale cerca di rispondere è dunque *quali siano i luoghi nei quali le aree industriali tendono a collocarsi?*

[37] *Sulla localizzazione delle industrie*. Weber 1909. Volume tradotto in lingua inglese nel 1929: Weber 1929.

[38] Lloyd, Dicken 1986, pp. 99-105; Tolman 1930, p. 110.

[39] Weber 1968, pp. 1-3.

[40] Fetter 1930, p. 233.

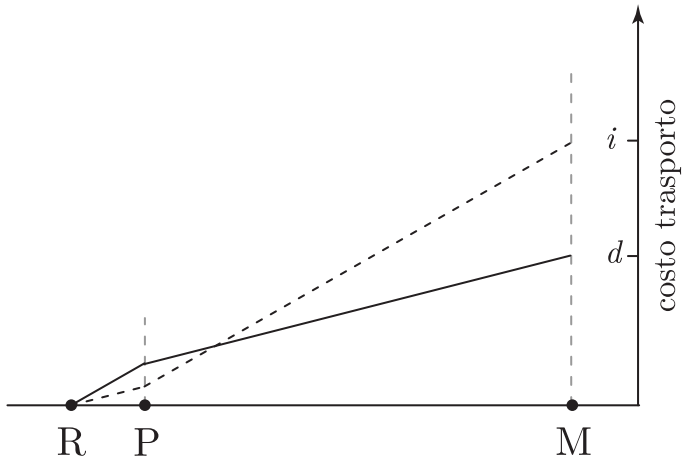


Fig. 3.10. Rappresentazione delle curve cumulative per aree industriali poste nelle vicinanze delle zone di estrazione delle materie prime. La curva tratteggiata (i) mostra il caso di un tipo di produzione che comporta un incremento dei pesi relativi tra materia prima e prodotti. La curva continua (d) mostra invece il caso del decremento dei pesi relativi. Come si pu osservare, $i > d$.

Il modello di Weber, oltre alle abituali schematizzazioni di tipo spaziale prevede dei requisiti o delle premesse relative a disponibilità e costo della manodopera che dovranno essere costanti nello spazio. Nello stesso modo, anche lo sviluppo tecnologico e le tecniche di produzione industriale sono considerate analoghe per un'intera regione.

Solitamente le aree di produzione industriale P, al fine di incrementare al massimo il profitto, si collocano nel punto dove portano a minimizzare i costi complessivi della produzione[41]. All'interno di una regione specifica, i costi delle materie prime, così come quelli della mano d'opera, dovrebbero essere considerati costanti. Saranno dunque ancora una volta i costi del trasporto a determinare la logica locazionale di questo tipo di attività.

In questo modello, oltre ai costi di trasporto, un ruolo chiave lo gioca la trasformazione del peso dei volumi nel passaggio da materia prima a prodotto industriale finito. Ogni processo industriale comporta la trasformazione delle materie prime. Questo processo può implicare, a seconda del tipo di produzione, una differenza di peso tra materia prima e prodotto finito. Così, la produzione di carta comporta una perdita di peso in relazione al legname originale necessario per produrla. Una tonnellata di carta non viene prodotta esattamente con una tonnellata di legno (polpa di cellulosa), ma con una quantità

[41] Merlin 1983, p. 166.

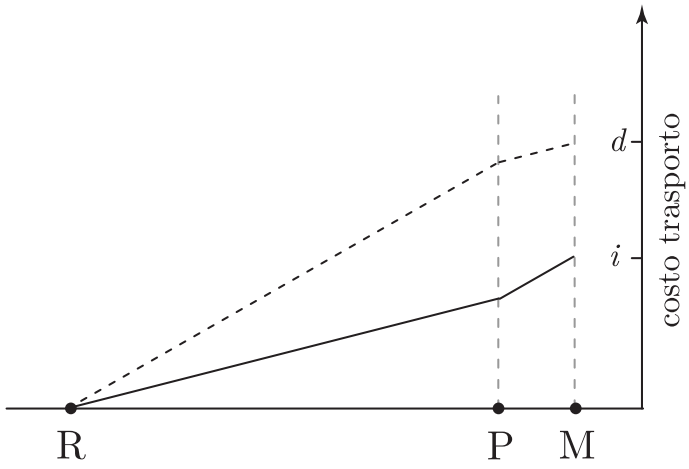


Fig. 3.11. Il grafico illustra i casi delle aree industriali poste nelle vicinanze dei mercati. La curva trattegiata (d) indica un decremento dei pesi relativi, mentre quella piena un incremento. In questo caso si osserva che $d > i$.

maggiore. Lo stesso può essere detto per l'argilla nel processo di produzione industriale dei mattoni[42].

Per comprendere la logica del modello weberiano prenderemo in considerazione un esempio ad una dimensione che prevede un unico luogo di estrazione delle materie prime e un unico mercato. Questo modello può essere facilmente convertito in un grafico di tipo cartesiano come quello illustrato nella figura 3.10[43]. Sulle ascisse viene rappresentata la distanza che separa l'area di estrazione R dal mercato M.

Il grafico illustra il caso di un'area industriale (P) posta nelle vicinanze dell'area d'estrazione. Nello schema vengono rappresentate due diverse situazioni. Una in cui, nel processo di trasformazione dei beni, si verifica un loro incremento (i, linea trattegiata)e una in cui, al contrario, si registra un loro decremento (d, linea continua) in relazione alle materie prime impiegate. Le due curve rappresentano i costi cumulativi di trasporto. Si può concludere dunque che, qualora il tipo di produzione comporti un incremento del peso dei beni prodotti in relazione alle materie prime utilizzate nella loro realizzazione, un luogo di lavorazione vicino alle aree di estrazione (R) risulterà meno vantaggioso che nel caso di tipi di produzione che prevedono invece un decremento del peso dei beni. La logica del modello porta dunque a concludere,

[42] Lloyd, Dicken 1986, p. 101.
[43] Haggett, Cliff, Frey 1977, p. 179, ed in particolare *fig.* 5.21.

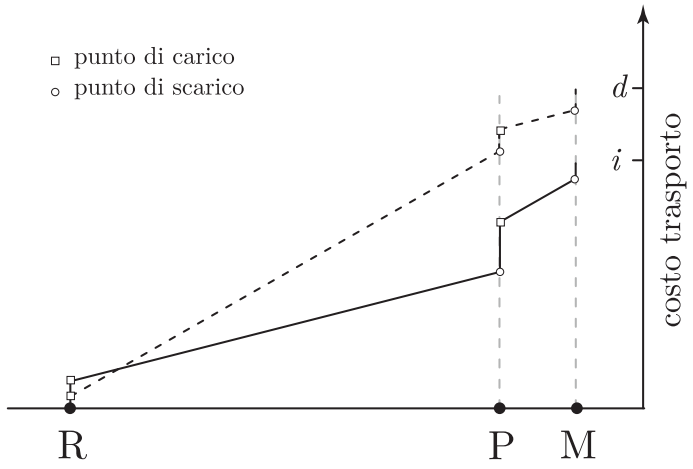


Fig. 3.12. Nell'applicazione della teoria di Weber è possibile includere anche altri elementi, quali spese di carico e scarico delle merci, per un più preciso computo dei costi di trasporto. Il grafico mostra come, in questo caso specifico, la differenza tra d e i sia meno marcata se si inseriscono i costi di carico e scarico.

ad esempio, come tipi di produzione quali carta o mattoni (ovvero produzioni che determinano una perdita di peso) vadano a collocarsi vicine alle aree di estrazione delle materie prime.

La differenza di inclinazione dei segmenti della curva è dovuta in sostanza alla diversità tra il costo di trasporto delle materie prime e il costo di trasporto dei prodotti. Una volta individuata una tipologia specifica, all'interno di uno spazio isotropico, il costo di trasporto per unità spaziale rimane costante. Ovvero, se x è il costo per il trasporto per $1km$ di $1m^3$ di argilla, $2x$ sarà l'equivalente per il trasporto della stessa merce per $2km$ e così via. I costi di trasporto crescono proporzionalmente all'incremento della distanza. Proprio per questo motivo, le due curve che rappresentano il costo del trasporto sono rette. È ovvio però che esso cambia a seconda dei beni. Infatti i costi di trasporto variano a seconda della tipologia del mezzo, del peso e del volume delle merci. Ad esempio, quelli su strada sono diversi per trasferire $1m^3$ di argilla o $1m^3$ di mattoni. Così come è diverso il trasporto per unità di volume di malto che di bottiglie di birra.

Nel grafico 3.11 invece è possibile osservare il caso di un impianto industriale posto nelle vicinanze del mercato. Sulle ordinate sono i simboli i e d utilizzati per indicare i costi cumulativi per il trasporto. Sempre in questo caso, i indica la curva cumulativa dei costi finali per il trasporto quando si verifichi un incremento del peso relativo, e d lo stesso valore ma in presenza di

un decremento. Il grafico permette di concludere come questo tipo di "posizione" sia più vantaggiosa, in termini di costo di trasporto, per tipi di produzione che comportano un incremento dei pesi dei prodotti in relazione alla materia prima. Come nel caso di un impianto di produzione e imbottigliamento della birra, ogni processo industriale che implica l'incremento del peso del bene prodotto in relazione alla materia prima, la localizzazione delle aree industriali risulta orientata verso il mercato M.

Se invece l'elaborazione delle materie prime comportasse una riduzione del loro peso, ovvero qualora il costo di trasporto di queste fosse superiore al costo di trasporto dei prodotti finali, allora le aree di produzione si orienterebbero verso quelle di estrazione. Nella figura 3.10, infatti, era già stata osservata la relazione $i > d$. Viceversa, qualora il tipo d'estrazione comporti un incremento del peso, l'area di produzione sarà orientata verso il mercato, come illustrato dalla figura 3.11, dove si verifica la relazione $d > i$.

Naturalmente, nel mondo reale, i prodotti dell'industrializzazione sono frutto di attività molto più complesse di quelle che possono essere descritte con un semplice modello a una dimensione. Anche solo per una matita bisogna tenere conto della provenienza del legno, della mina in grafite, della vernice esterna, della colla, dell'eventuale gomma e del suo supporto in metallo e dell'imballaggio, quest'ultimo a sua volta un prodotto industriale.

Il modello originale di Weber si riferisce dunque alla localizzazione del punto ideale su di un piano. Nel caso della figura 3.13 sono previste due fonti di materie prime R_1 e R_2, un mercato M e un'area di processamento P. In questo caso il modello dovrebbe aiutare a capire quale sia sulla base dei parametri di costo, la località ideale in cui collocare l'industria.

Potremmo dunque immaginare come il costo complessivo della produzione (esclusa la mano d'opera) debba tenere conto del trasporto delle singole materie prime all'impianto industriale e, infine, dall'impianto al mercato. Nello specifico:

- $r_1\overline{R_1P}$: costo totale del trasporto delle materie prime dalla prima fonte all'area industriale;
- $r_2\overline{R_2P}$: costo totale del trasporto delle materie prime dalla seconda fonte all'area industriale.
- $m\overline{MP}$: costo totale del trasporto dei prodotti finiti dall'impianto industriale al mercato;

Naturalmente, nel calcolo dei singoli costi si deve tenere conto anche delle differenze di peso, ovvero degli esiti della trasformazione delle materie prime nel processo industriale. Sulla base di questo principio, la posizione ideale dell'area industriale sarà dunque quella dove si otterrà il risultato minimo dall'equazione:

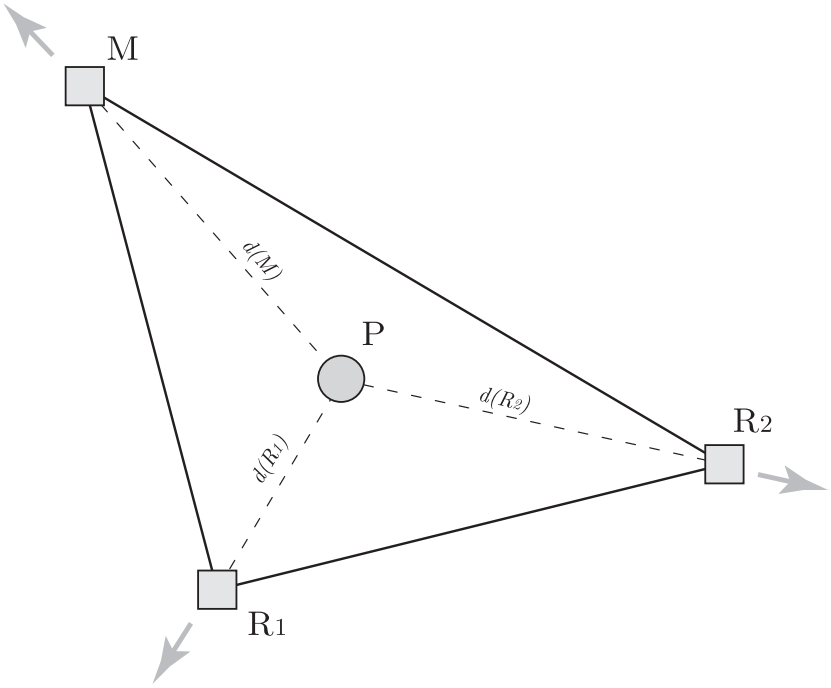


Fig. 3.13. La produzione industriale raramente si fonda sul processamento di un'unica materia prima. Perciò la teoria locazionale di Weber si riferisce più a processi che possono essere compresi nel piano che non su modelli astratti ad un'unica dimensione.

$$w_1 r_1 \overline{R_1 P} + w_2 r_2 \overline{R_2 P} + m\overline{MP}$$

dove w_1 e w_2 rappresentano i pesi complessivi necessari per la produzione di un'unità di peso di merce già pronta, r_1 e r_2 i costi relativi al trasporto per chilometro delle materie prime e m il costo per chilometro dei prodotti finiti.

La posizione ideale per il punto P è dunque quella che consente di ottenere per l'equazione precedente il risultato più basso. Naturalmente il calcolo della posizione ideale non si può definire con il semplice computo del baricentro ideale. Questo può funzionare solo quanto i costi di trasporto (r_1, r_2 e m) si equivalgono; ovvero quando il costo per il trasporto per ogni chilometro di una tonnellata di prodotti finiti equivale a quello di una tonnellata di materie prime del primo tipo e una tonnellata di quelle del secondo tipo.

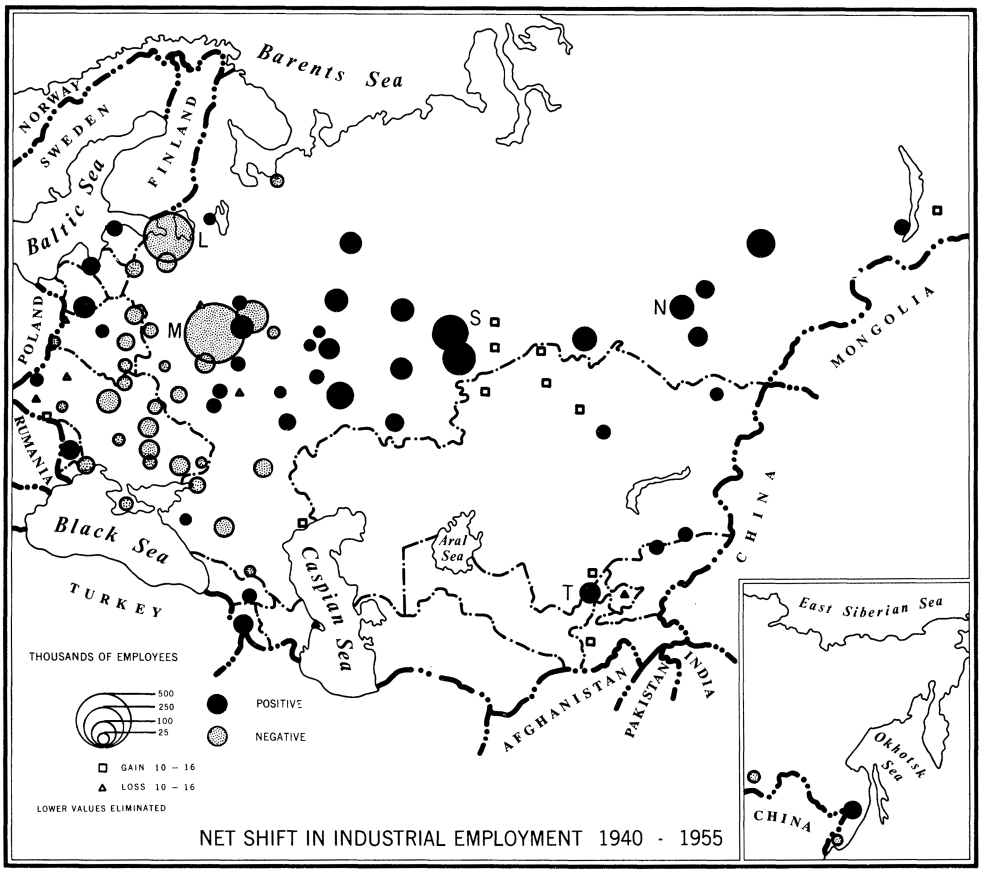


Fig. 3.14. L'applicazione del modello di Weber nella realtà deve tenere conto non solo della distribuzione di materie prime ed energia, ma anche della distribuzione spaziale della mano d'opera. Molte ricerche che hanno sfruttato la logica della teoria della localizzazione delle industrie si sono concentrate proprio sulla distribuzione della popolazione, le sue trasformazioni e l'impatto nella distribuzione spaziale dei complessi industriali (Rodgers 1974, p. 234).

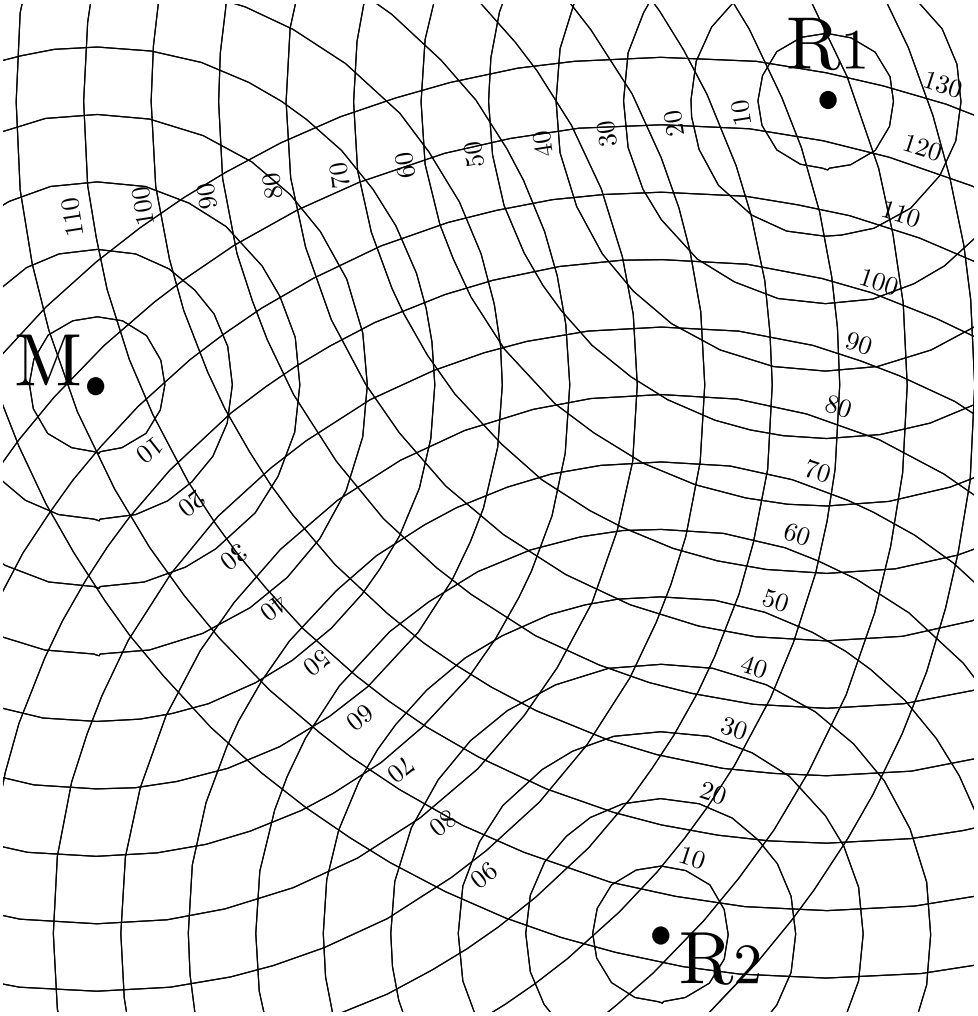


Fig. 3.15. Il primo passo nel processo di identificazione in un piano dell'area ideale per una zona industriale è quello di tracciare la rappresentazione grafica dei costi di trasporto progressivi a partire dalle aree di estrazione o dal luogo di mercato. Questi cerchi concentrici vengono normalmente definiti *isotime*.

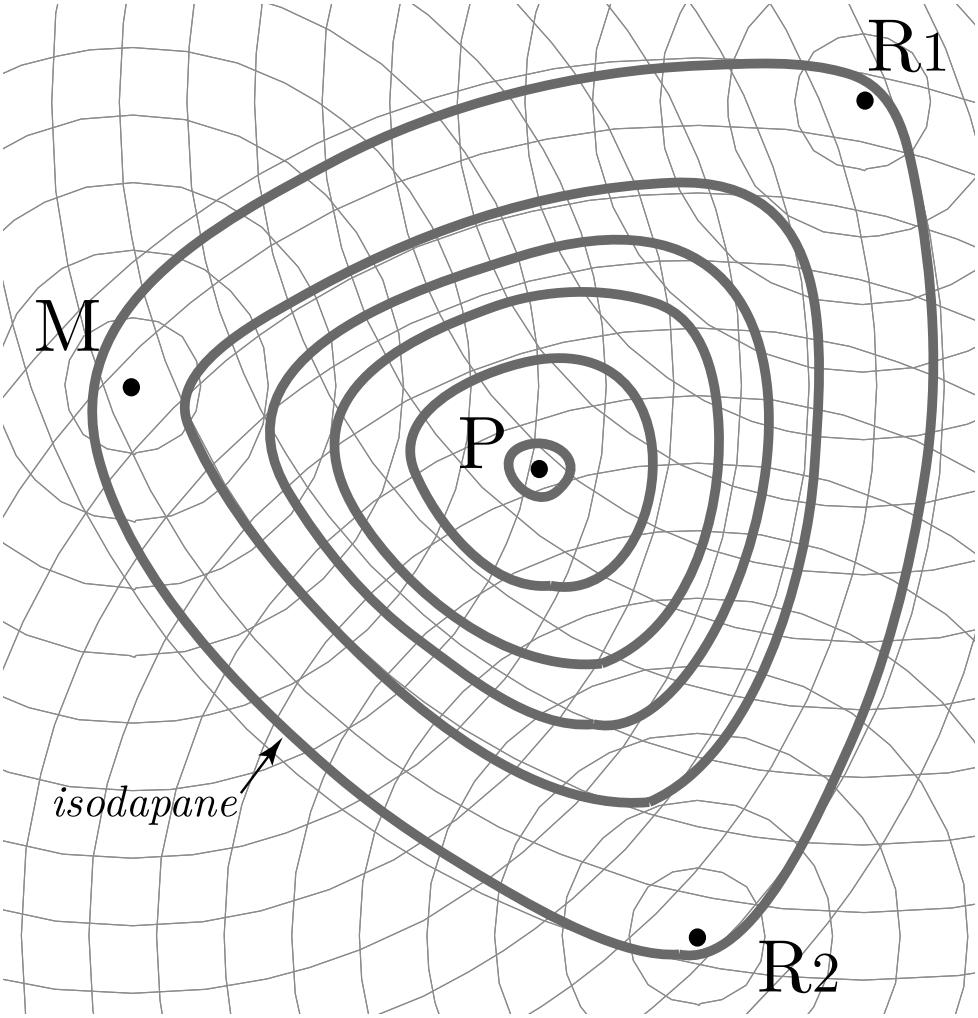


Fig. 3.16. Successivamente, a partire dalle *isotime*, si procede al computo della somma dei costi complessivi di trasporto di materie prime e merci da ogni punto del piano. Si passa poi a unire con delle linee i luoghi in cui tale somma o costo complessivo di trasporto è equivalente. Queste linee a costo equivalente vengono definite isodapane. Solitamente l'isodapana al centro segna il punto di minore costo.

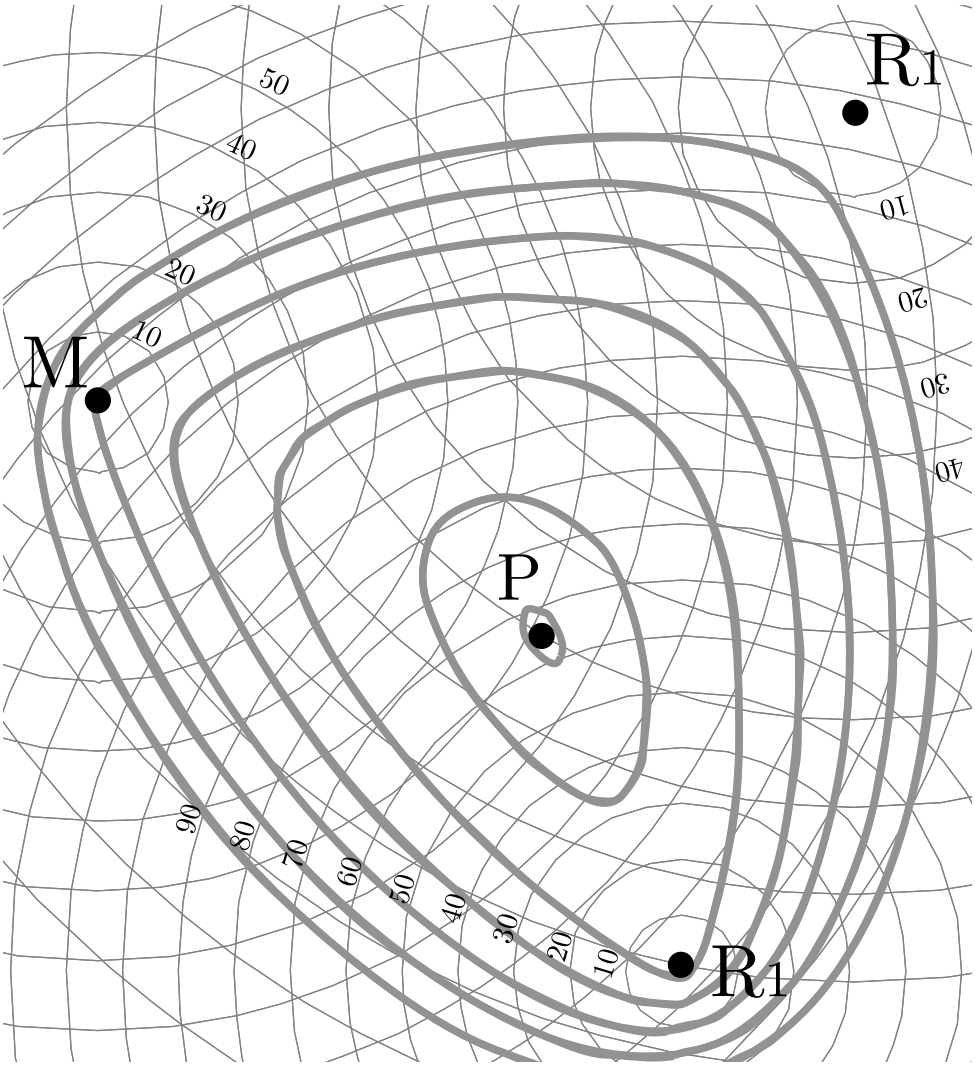


Fig. 3.17. La variazione dei costi di trasporto da e verso uno dei punti interessati comporta la distorsione delle isodapane.

Si tenga presente dunque che il modello di Weber deve tenere conto di tre tipi di variabili: distanze, pesi e costi di trasporto. Dunque l'identificazione dell'area P risulta particolarmente complessa. A risolvere questo tipo di quesito si sono proposte nel tempo diverse soluzioni, tra le quali quella che gode di maggiore popolarità proviene dallo stesso Weber. Si tratta di una soluzione empirica che consente, semplicemente attraverso la rappresentazione schematica dei costi, di identificare graficamente il punto ideale.

In primo luogo si procede alla rappresentazione intorno ai punti M, R_1 e R_2 del prodotto tra distanza e costi di trasporto per tipologia di bene con l'aiuto di una serie di cerchi concentrici, ognuno dei quali avrà un intervallo spaziale in relazione al suo valore specifico. Il risultato è dunque un gioco di grafici dove tra le altre cose è molto facile perdersi data la natura intricata del risultato. Successivamente si procede a congiungere i punti per i quali i costi complessivi della loro somma sono uguali (vedi figure da 3.15 a 3.17).

3.3 Il modello delle località centrali

Osservando l'organizzazione delle forme di stanziamento umano si può notare come i centri, presi singolarmente, presentino diverse forme e dimensioni. Viaggiando in macchina per le campagne si possono attraversare villaggi di diversa consistenza, vedere in lontananza gruppi o case rurali sparse. Guardando le carte si può notare che le località piccole e grandi sono distribuite in sequenze alternate, a volte regolari, lungo gli assi stradali[44].

L'insediamento umano si articola in una complessa struttura composta da una serie di gerarchie insediative che sembrano lasciare intuire una qualche forma di ordine spaziale. Uno dei primi studiosi a formulare un modello sulla formazione e sulle funzioni delle gerarchie e sull'organizzazione spaziale dell'insediamento è stato il geografo tedesco Walter Christaller che nel 1933 pubblicava la sua tesi di dottorato *Die zentralen Orte in Süddeutschland*[45] – titolo dal quale deriva il nome del modello – nella quale egli esponeva in modo dettagliato i principi alla base della sua teoria.

La figura di Walter Christaller è fra le più controverse della disciplina geografica. La sua teoria era ben nota negli Stati Uniti prima di ricevere qualche forma di riconoscimento in patria. Nessun altro contributo è stato così criticato, ma allo stesso tempo così influente nelle diverse scuole geografiche durante il XX secolo[46]. Lo dimostra l'esteso numero di saggi e articoli aventi come

[44] Christaller 1980, pp. 25-26.

[45] Christaller 1968. Traduzione in italiano: Christaller 1980.

[46] Per cenni biografici su Walter Christaller consultare Hottes 1983, pp. 81-82; Berry, Harris 1970, pp. 116-119;

soggetto, questo modello. Il lavoro di Christaller arrivò a godere di una certa popolarità molto tardi rispetto alla prima pubblicazione della sua tesi.

I punti che il suo modello vuole trattare sono descritti con chiarezza dall'autore stesso quando afferma: «Ma come possiamo trovare una spiegazione generale per le dimensioni, il numero e la distribuzione delle città? E come possiamo scoprire le leggi che le regolano? Può forse la ricerca geografica fornire una spiegazione?»[47]. Lo scopo del suo lavoro era sostanzialmente identificare i meccanismi attraverso i quali, all'interno di un processo storico, certi centri acquisissero un'importanza maggiore di altri[48].

Allievo di Alfred Weber, il suo modello era costruito e impregnato della logica e dei fattori delle variabili tipiche della geografia funzionalista: distanza, costo, beni, offerta, profitto ecc[49].

Al di là dei contributi empirici che può avere realizzato Christaller (soprattutto aprendo la strada a discendenti intellettuali) forse l'apporto più importante è stato quello di avere fornito al dibattito sul pensiero geografico il senso stesso e le chiavi interpretative sullo spazio; soprattutto se analizzate sotto l'ottica degli studi regionali. Nella volontà di identificare delle spiegazioni alle maglie di stanziamento, lo studioso attribuisce allo spazio e alle sue dinamiche un ruolo chiave nei processi di definizione delle maglie di stanziamento: «Solo pochi autori di trattati economici sono ben consapevoli del fatto che tutti i rapporti ed i processi economici sono condizionati, senza alcuna eccezione, dallo spazio e che quindi non possono nemmeno essere concepiti senza il riferimento spaziale, che è un loro elemento costitutivo»[50]. Sia ben chiaro: le chiavi interpretative di cui si diceva sopra, per quanto elementari, sono efficaci. Tanto che permetteranno ad altri studiosi di ampliarle e strutturarle all'interno di modelli più complessi.

Nel tracciare la sua tesi, Christaller non si affida al concetto di popolazione, ma introduce l'idea di "importanza della località" intesa come centralità geografica[51]. Il ruolo di un centro abitato non dipende (quantomeno direttamente) dal numero di residenti, bensì dalla portata delle dinamiche che scaturicono al suo interno: «L'"importanza" non è una grandezza matematica o la semplice somma degli abitanti, [...] bensì il prodotto dell'azione comune della popolazione. A questa azione comune da noi detta "importanza", ci si richiama ogni qualvolta si definisce una città "viva", "fiorente", o "importante".»[52]

[47] Christaller 1980, p. 27.
[48] Merlin 1983, p. 189.
[49] Tinacci Mossello 1990, p. 152.
[50] Christaller 1980, p. 33.
[51] Un'analisi sul ruolo della centralità nell'opera di Christaller si trova in Preston 1971, pp. 136-138; Dacey 1965, pp. 111-115; Kenyon 1967, pp. 736-739.
[52] Christaller 1980, p. 45.

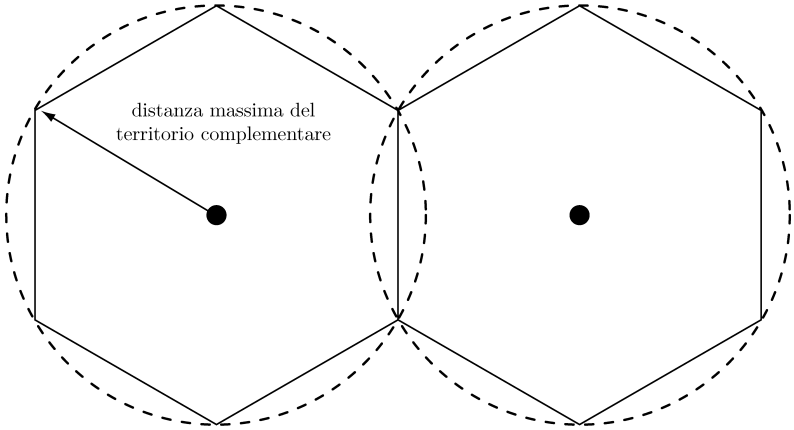


Fig. 3.18. I "territori complementari" di Christaller sono definiti dal prezzo massimo che può essere pagato per un servizio. In effetti, al costo del servizio si deve aggiungere il costo del trasporto della persona fino alla località centrale. Il territorio complementare rappresenta l'area geografica compresa entro i limiti determinati dal prezzo massimo per un determinato servizio. All'interno della logica della domanda e dell'offerta, lì dove finisce il territorio complementare massimo per un certo servizio esiste spazio per una nuova località centrale.

Secondo Christaller non sono le località ad essere importanti quanto i "servizi" prestati al loro interno. Con il concetto di servizio egli introduce un concetto amplio: un barbiere, un dentista, un cinema, un ospedale, una scuola media, un'università, la piazza del mercato ecc. Attività che possono essere svolte all'interno di insediamenti di una certa consistenza e non sparse nel territorio. Sempre secondo Christaller, l'importanza, o quello che normalmente si considera importanza, deriva dal servizio (o dai servizi) e non dalla dimensione della località e dal numero dei suoi abitanti. Ad esempio, una città piccola può essere di grande importanza perchè sede di una università.

Riallacciandosi ad una visione weberiana della geografia, egli introduce il concetto di "territorio complementare" o "portata" del servizio. Si tratta semplicemente del limite spaziale, espresso sotto forma di distanza, entro il quale lo sfruttamento di un servizio è praticabile[53]. Occorre ricordare che l'abitante delle aree rurali, al costo del servizio, per esempio un dentista, deve sommare anche quello del trasporto per giungere alla località centrale. Si tratta di un equilibrio tra convenienza, concorrenza e profitto, mediati dai costi di traspor-

[53] Il concetto di territorio complementare verrà ampliato successivamente da Lösch (paragrafo 3.4).

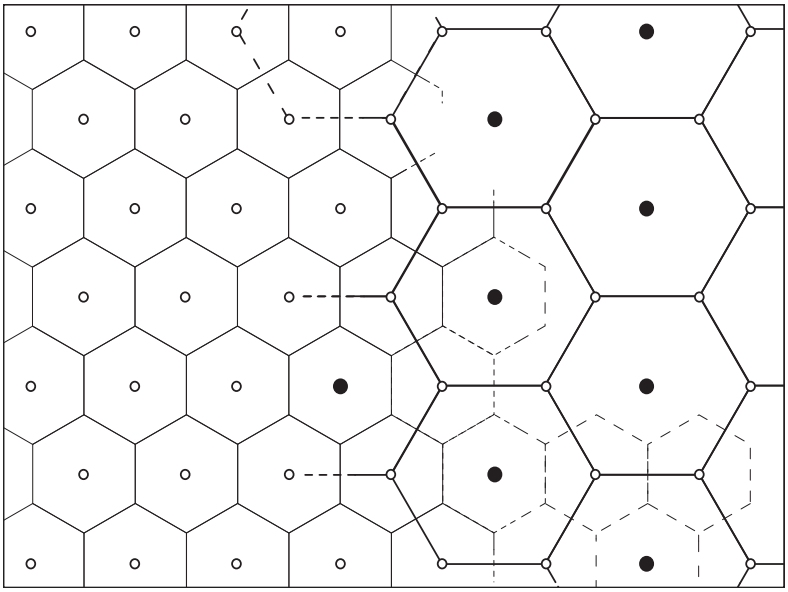


Fig. 3.19. La maglia dei territori complementari relativa ad un servizio specifico acquista una forma ad alveare. Le località di rango inferiore (punti bianchi) sono poste ai vertici dei poligoni di Thiessen delle località di rango superiore (punti neri). I vertici sono caratterizzati da proprietà particolari dato che le località inferiori in questi luoghi possono usufruire dei territori di pertinenza e dunque dei servizi di tre località di rango superiore. Inoltre si osservi come i punti di gerarchie inferiori e superiori facciano parte di un'unica trama triangolare. In altre parole, la geometria christalleriana fa sì che insediamenti di gerarchia superiore siano anch'essi di gerarchia inferiore.

to. In questo senso, Christaller si riallaccia alle basi dello spazio economico di von Thünen.

Dato che la popolazione costituisce un bacino potenziale per ogni tipo di servizio, e questi devono essere sostanzialmente garantiti all'interno di una logica di mercato, le località centrali (cioè quelle che forniscono il servizio) devono essere distribuite in modo omogeneo sull'intero territorio. Non vi possono e non vi devono essere aree non coperte da offerte di servizio accessibili e praticabili in termini di costo; ovvero il prezzo del servizio più il costo di percorrenza dalla residenza alla località centrale. Nella strutturazione naturale dello spazio sociale, tale processo avviene grazie a "territori complementari" che per la loro conformazione finiscono per sovrapporsi, come illustrato nella figura 3.18. I loro confini vengono tracciati seguendo la logica dei poligoni di Thiessen. Per questo motivo, più località centrali sorgono ad intervalli regola-

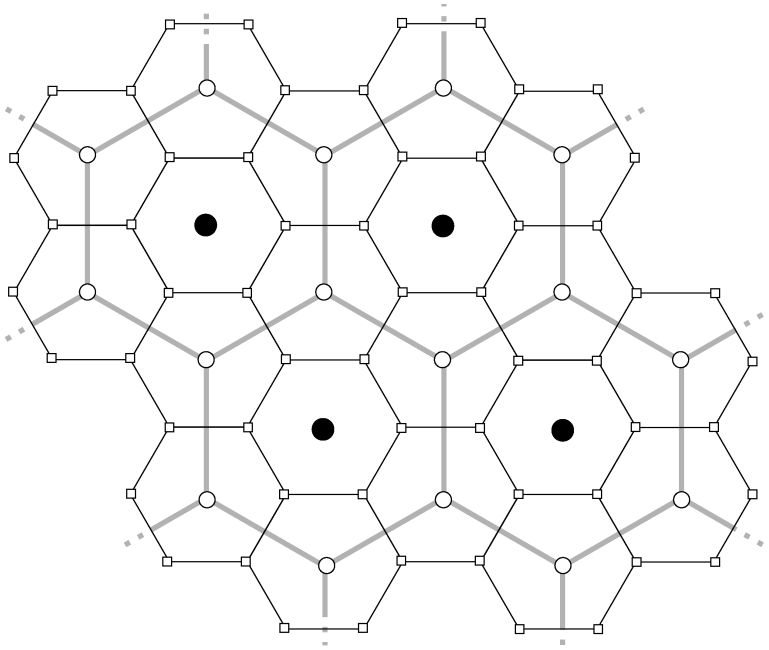


Fig. 3.20. Le località di rango inferiore (cerchi bianchi) replicano, a livello gerarchico, lo schema descritto nella figura 3.19. Infatti, sui vertici di questo livello si trovano insediamenti appartenenti ad un rango ancora inferiore (quadrati piccoli bianchi). Un altro aspetto degno di nota è che nel grafico sono ben visibili tre ranghi insediativi, inseriti all'interno di un'unica griglia triangolare regolare.

ri nello spazio. L'intero modello christalleriano è basato su una distribuzione degli insediamenti su di una maglia a triangoli equilateri, che costituisce l'organizzazione ideale nello spazio. I poligoni di Thiessen dei punti che formano questa trama presentano dunque la caratteristica forma ad esagoni. Naturalmente questa organizzazione regolare risponde a requisiti tipici di uno spazio astratto di tipo euclideo ed isotropico.

La struttura fisica del modello di Christaller è legata infatti allo schema dei poligoni di Thiessen (vedi paragrafo 3.1). Nella parte destra della figura 3.19 si possono osservare una serie di località centrali e la distribuzione delle aree di pertinenza. Ai vertici di queste aree si trovano altri insediamenti che, come si può osservare, appaiono di rango inferiore. Va sottolineato, per maggiore chiarezza, che tale rango deriva sempre del concetto di servizio offerto.

Gli insediamenti di rango inferiore non sono prodotto dei centri superiori. I centri dei diversi ranghi appartengono ad un'unica rete originaria. Va chiarito però che livello inferiore e superiore non si trovano in assoluto su due dimen-

sioni spaziali diverse. Nella figura 3.19 si può osservare come località di rango superiore e rango inferiore si alternino all'interno di un'unica trama di territori complementari più piccoli, rilevate da una maglia di poligoni di Thiessen girata di 90 gradi. A differenza di quanto si potrebbe pensare, dunque, ranghi inferiori non sono il prodotto della scelta di luoghi strategici, ma ranghi superiori ed inferiori si integrano all'interno di una maglia estesa e primordiale che nel tempo determina, attraverso processi di evoluzione dei servizi, la nascita di una gerarchicamente superiore. Ed è perciò importante ricordare sempre che la matrice del modello parte dal basso e non dall'alto. Se si osserva la figura 3.19 da sinistra verso destra, si può notare come i centri di rango superiore condividano le posizioni di centri appartenenti anche ai ranghi inferiori.

Sempre secondo Christaller, certi tipi di servizio andranno a concentrarsi su certe località. Questo finirà per assegnare a quelle località un peso ancora maggiore e di conseguenza un territorio complementare più vasto. Nello stesso tempo, il fatto che due tipi di servizio siano concomitanti in un determinato territorio, non significa che le aree dei suoi territori complementari coincidano. Prendiamo l'esempio del sistema scolastico. Traducendolo alla logica del modello delle località centrali, possiamo contare almeno quattro tipi di servizi: scuole materne, elementari, medie-superiori e università. Prendendo in considerazione la funzione della scuola materna, e le necessità proprie dei bambini, la attenzioni specifiche per il singolo, ed in particolare la vicinanza ai propri genitori, la loro distribuzione deve essere capillare. Avanzando nella scala di servizi, e soprattutto superando il limite delle scuole dell'obbligo, se da una parte si assiste alla diminuzione progressiva della popolazione, parallelamente, man mano che crescono, gli studenti diventano più indipendenti. Inoltre, il servizio, crescendo il suo livello in relazione all'età degli utenti, diventa più specializzato. Non sarebbe fattibile in termini economici una rete fitta di università. La matrice del modello viene definita dall'acquisizione progressiva e ordinata nel tempo di funzione di servizi su scale sempre maggiori da parte di insediamenti particolari.

Occorre sottolineare però che non esistono località specializzate esclusivamente nel servizio universitario. Le città con università possiedono anche scuole superiori, elementari ecc. Dunque non è il centro a determinare la portata della propria area complementare, bensì il servizio offerto.

La sequenza e gli intervalli secondo la maglia gerarchica dei servizi, per Christaller costituiscono la chiave per capire le motivazioni della distinzione dei paesaggi insediativi.

Come è stato già affermato prima, la teoria del luogo centrale fu oggetto di molte critiche. La principale era quella sulla natura statica della trama christalleriana. Benché l'opera di Christaller presentasse diversi esempi di ap-

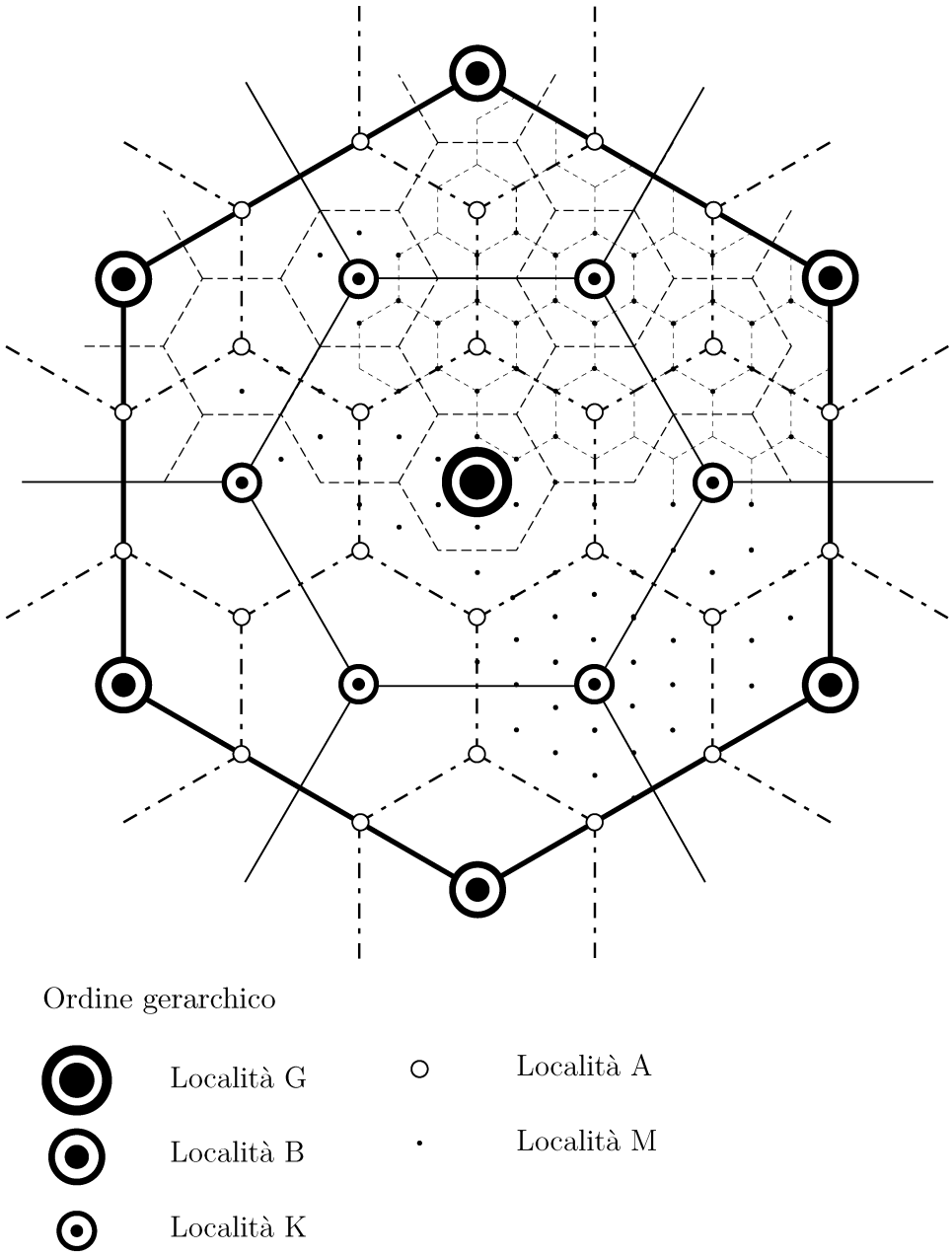


Fig. 3.21. Lo schema descritto dalla figura 3.20 si ripete su una struttura gerarchica organizzata su più livelli. Questo processo comporta un progressivo ampliamento delle aree complementari. Servizi appartenenti a ranghi superiori vanno così a coprire territori più vasti.

plicazione e verifica del suo modello con la realtà distributiva di diverse regioni della Germania meridionale (Monaco, Norimberga, Stoccarda, Strasburgo e Francoforte), il suo modello non permetteva di raggiungere interpretazioni soddisfacenti della complessità distributiva.

Inoltre, pur avendo Christaller stesso introdotto la distinzione dei servizi, e la diversità della logica spaziale delle trame, la teoria del luogo centrale appare troppo rigida per l'interpretazione di realtà al di fuori dei contesti propri nei quali la sua teoria si è originata.

Fu lo stesso Christaller a definire la distinzione degli effetti a seconda della tipologia dei servizi. In particolare identificava tre categorie fondamentali di servizio: mercati (o scambi commerciali), comunicazione (o trasporto) e amministrazione. Secondo Christaller, questi tre principi, che rappresentano categorie specifiche dei beni e servizi, nella realtà diventano incompatibili in termini spaziali o topologici. Ovvero, l'organizzazione dello spazio con una sovrapposizione delle tre tipologie di reti fa sì che i centri dei poligoni su una sequenza gerarchica non andassero poi effettivamente a coincidere[54].

La logica delle distinzioni sarà la chiave attraverso la quale alcuni dei suoi successori potranno sviluppare modelli più complessi per colmare le lacune ancora presenti nella teoria del luogo centrale.

3.4 Il modello di Lösch

August Lösch (1906-1945) forse rappresenta la figura meno nota ma allo stesso più importante della scienza locazionale in ambito geografico. Scomparso in giovane età sul finire della seconda guerra mondiale, seppe far evolvere e trasformare definitivamente la logica della geografia economica grazie alla pubblicazione postuma del volume *Die räumliche Ordnung der Wirtschaft*[55].

L'opera di Lösch non si limita all'esposizione della sua teoria sulla struttura gerarchica regionale, ma si spinge ben oltre, dando forza e novità a diverse correnti, incluse quelle di Weber e Von Thünen, cercando, per così dire, una sorta di compendio generale della logica applicativa della teoria locazionale in ambito economico[56]. Il suo lavoro resta uno dei principali punti di riferimento nel settore degli studi locazionali; la diffusione dell'opera avvenne durante gli anni Cinquanta, specialmente nei paesi anglosassoni solo dopo che fu tradotta in lingua inglese[57].

[54] Carter 1975, pp. 59-64.
[55] *L'ordinamento spaziale in economia*: Lösch 1940.
[56] Dunn 1955, p. 367.
[57] Parr 1973, p. 187.

L'aspetto più noto (ma sicuramente non l'unico) del lavoro è quello che per comodità chiameremo il modello di Lösch. Benché basato sui principi della logica gerarchica di Christaller, lo schema di Lösch introduce nuovi elementi che non solo integrano, ma trasformano totalmente la logica del modello del suo predecessore. Si può dire che il nuovo paradigma, benché non pienamente esplicitato, acquisti l'articolazione che lo rende ancora più appropriato alla descrizione della complessità delle forme dell'insediamento umano[58].

L'elemento rivoluzionario dell'opera di Lösch, a differenza di quelle dei suoi predecessori, sta nell'aver fornito un modello basato non sulla definizione statica di un paesaggio geografico, bensì nell'aver proposto uno schema spaziale dove interrelazioni a livello di produzione, consumo, collegamenti, scambio, trasporto, amministrazione e governo si integrano in modo coerente all'interno di un sistema[59].

Il paesaggio löschiano si basa in sostanza sulla sistemazione formale dei principi di organizzazione all'interno del modello christalleriano. In primo luogo egli definisce con maggiore chiarezza il concetto di territorio complementare ampliandone con grande precisione e rigore le implicazioni del trasporto nella loro definizione[60]. Egli chiarisce così il concetto di soglia di mercato, cioè il limite oltre il quale un bene non è commercializzabile o un servizio non è praticabile, dato che i costi di trasporto comporterebbero un annullamento dei profitti sulle vendite o un dispendio eccessivo per i fruitori. I territori complementari di Christaller assumono così la forma di un cono d'offerta. L'ambito di diffusione di un bene esterno definito dalla domanda o dall'offerta rappresenta una soglia oltre la quale si entra nell'ombra di un nuovo cono d'offerta.

Sempre basato sul piano isotropico uniforme a fertilità omogenea, il modello locazionale di Lösch si basa sull'identificazione della misura K per ogni singolo insediamento. In questo modello, il paesaggio originale è anch'esso composto da una rete omogenea di insediamenti (simile al paesaggio christalleriano) che, nel corso della sua storia, tende a sviluppare un'organizzazione che porta alla differenziazione articolata – o eterogenea – delle gerarchie a livello spaziale. La misura K non è altro che il numero di insediamenti serviti da un determinato centro. In altri termini, K rappresenta la portata massima di un territorio complementare.

Nella figura 3.22 è possibile osservare, nel caso a (logica di mercato), come K sia uguale a 3. Questo perché il territorio complementare minimo o di base serve sostanzialmente 3 insediamenti. Nello specifico $K = 1 + (6 \times \frac{1}{3}) = 3$. Ov-

[58] Eaton, Lipsey 1976, p. 77.

[59] Stolper 1959.

[60] L'integrazione del concetto di territorio complementare era già stato integrato dallo stesso Lösch nel saggio *The nature of economic regions*: Lösch 1938, pp. 71-72.

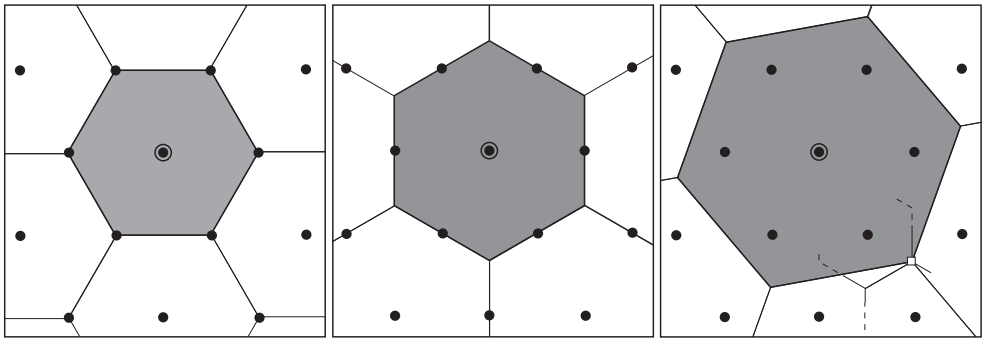


Fig. 3.22. Secondo Christaller esistevano tre schemi gerarchici per le regioni complementari: mercato, comunicazione e amministrazione. Lösch introduce un valore o indice di riferimento per tali tipologie. Questo valore, solitamente noto come numero K, rappresenta il numero di insediamenti serviti da una determinata area complementare. L'immagine, che presenta gli schemi elementari per ognuna di queste tipologie, è basata sulle figure 24-26 in Lösch 1959, p. 117.

vero, l'area complementare minima serve l'insediamento stesso in cui si trova il mercato, oltre a un terzo dei sei insediamenti posti ai vertici dell'esagono. Così $K = 3$ rappresenta l'area complementare, base (minima) di mercato.

La razionalizzazione della logica del traffico o delle comunicazioni prevede una diversa rotazione dei territori complementari. Si osservi anche la rotazione che comporta, fra l'altro, un incremento dell'area e della portata della superficie dei territori complementari. In questo senso, il centro inteso come punto di supporto per il traffico di merci o persone va a offrire il proprio servizio a 4 insediamenti: $K = 1 + (6 \times \frac{1}{2}) = 4$. Ovvero, a sé stesso oltre che a metà dei 6 siti posti ai lati del poligono equivalente al proprio territorio complementare.

In questo modello viene introdotto un nuovo principio di rotazione per l'identificazione più efficiente della logica amministrativa[61]. Il territorio complementare elementare per l'amministrazione secondo Lösch doveva avere una rotazione come quella rappresentata nel caso c della figura 3.22, dove il vertice di quest'area corrisponde non a un sito, ma a uno dei vertici dei poligoni di Thiessen della maglia originaria, come indicato dai piccoli poligoni tratteggiati in basso a destra. Con la logica amministrativa di base prevede $K = 1 + (6 \times 1) = 7$.

[61] Vedi Parr 1973, p. 186; Haggett, Cliff, Frey 1977, pp. 148-149.

I valori di K possono estendersi all'infinito. Nella figura 3.23 si può osservare come le trame crescenti possano effettivamente coprire un intero territorio.

rango n.	**K**
1	3
2	4
3	7
4	8
5	12
6	13
7	16
8	19
9	21
10	25
11	27

Tabella 3.1. Nel suo trattato Lösch non fa riferimento al valore K bensì ad un rango progressivo (*Area n.* Così nella sua esposizione quando si fa riferimento a località di rango 1 ci si riferisce a centri con un territorio complementare equivalente a K=3; 2 equivale a K=4 e così via. LÖSCH 1959, p. 119.

Continuando con queste tre possibili combinazioni di rotazioni, si può giungere a costruire una sequenza crescente in cui nuove reti K vengono definite con un maggiore numero di località servite. Ogni categoria di servizio sarà rappresentato da un valore specifico K. Così, ad esempio, le scuole medie potrebbero essere rappresentate da $K = 7$, le superiori da $K = 13$ e le università da $K = 21$. Ogni valore di K semplicemente definisce la rotazione e la dimensione della trama di esagoni.

Questo modello, che prevede una rotazione costante e crescente della rete di esagoni, poneva il problema di come far combaciare i sistemi all'interno di un unico schema condiviso da tutte le possibili K. La soluzione proposta da Lösch andava nella direzione del concetto di località centrale, rafforzandone ancora di più il concetto.

Il cosiddetto paesaggio löschiano trae origine da un unico sito centrale. In questo sito andranno a sovrapporsi sostanzialmente $K = 3$, $K = 4$, $K = 7$, $K = 9$ e così di seguito. Questo perché, all'interno della logica del modello, questa località centrale offrirà tutti i tipi di servizi al territorio circostante. In altre parole, la località centrale costituisce l'origine di tutte le trame K.

Il modello di Lösch è schematizzato dalla figura 3.23 nella quale si può osservare la progressione di poligoni dai più piccoli a quelli di maggiore esten-

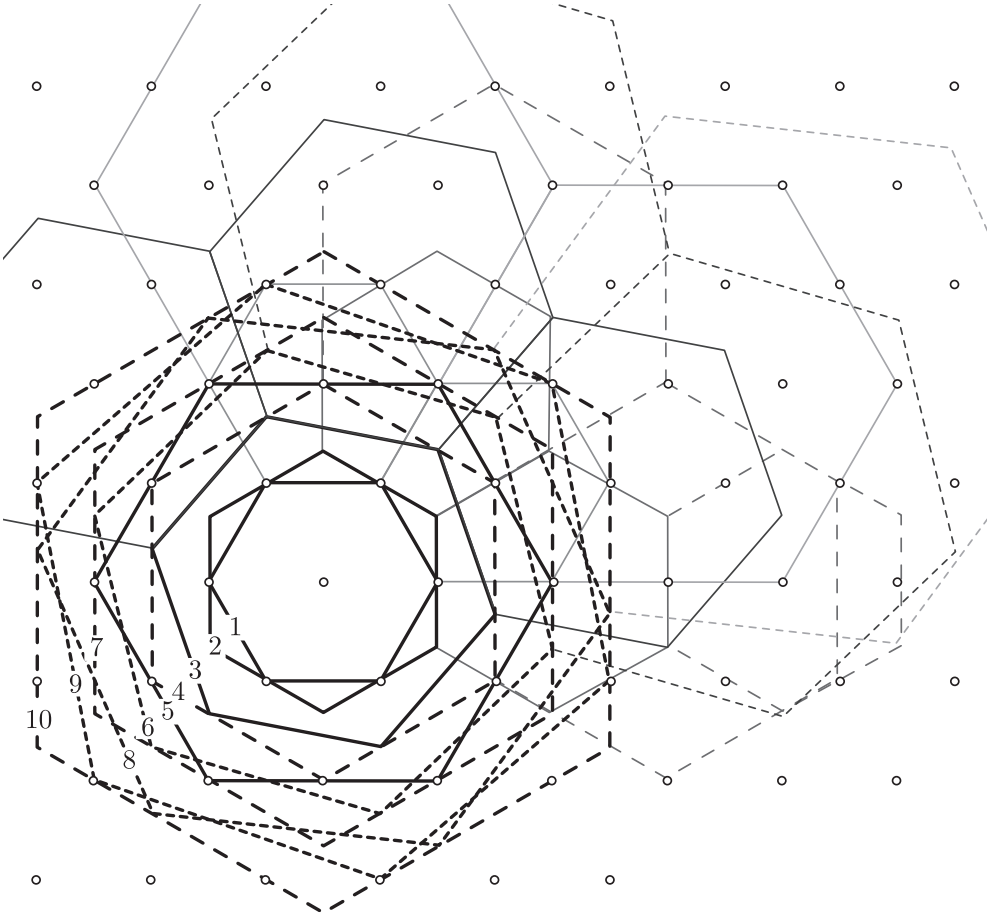


Fig. 3.23. Il paesaggio löschiano.

sione. Come si può notare, questi vanno a generare una trama, fitta di possibili combinazioni. Ognuno dei 10 poligoni raffigurati rappresenta solo il generatore di una rete che si estende su tutta la regione. In altre parole, l'immagine mostra come sia un unico centro ad acquistare la funzione di località centrale, sempre secondo la logica di Christaller. Occorre notare che Lösch, proprio per la complessità dei grafici che ne derivano, usa semplici numeri per chimarare le diverse gerarchie K. Così 1 per $K = 3$, 2 per $K = 4$, 3 per $K = 7$ (vedi tabella 3.1).

La rotazione delle reti K comporta però una dinamica nuova nell'organizzazione dell'insediamento. Occorre ricordare che, nella sovrapposizione dei diversi livelli gerarchici, il centro dei poligoni di ogni tipo di rete cadrà sempre

in corrispondenza di un centro della trama regolare sottostante. In altre parole, il centro di qualsiasi possibile poligono del paesaggio löschiano sarà sempre un sito. Di fatto, quando si osservano i siti di un determinato territorio possiamo notare che non vi è una regolarità nella coincidenza dei servizi offerti; in altri termini, non tutti i siti appartenenti ad un'unica categoria offrono gli stessi servizi. Non tutti i centri di medie dimensioni possiedono un cinema; e non tutti possiedono una centrale di polizia. Ce ne sono alcuni in cui coincidono i due servizi, ma altri dove ce n'è solo uno dei due, oppure nessuno.

Estendendo gli schemi K, Lösch fu in grado di ottenere la figura 3.24 che rappresenta un quadrante di un grande territorio di forma esagonale. Il grafico mostra, per ogni punto, una serie di numeri. Questi corrispondono, secondo la tabella 3.1, alla gerarchia K. Così, un insediamento che presenta i numeri 1245 indica un sito che costituisce il centro di un poligono $K = 3$, $K = 4$, $K = 8$ e $K = 12$. La località centrale (generatore) porta i numeri $1 - 10$; il che sta ad indicare che questo insediamento, come località centrale, fornisce tutti i servizi previsti.

Il risultato è un paesaggio di grande complessità, ben diverso dalla natura statica di quello christalleriano. Di fatto Christaller aveva identificato alcune delle differenziazioni a livello spaziale dell'organizzazione del paesaggio a seconda della tipologia dei servizi; solo che non era riuscito ad integrarle in un unico sistema. A differenza di Christaller, l'elemento chiave nel modello di Lösch, è l'enorme differenziazione delle funzioni e dei servizi degli insediamenti della rete di punti sottostante. La nuova struttura gerarchica determina funzioni diversificate. Il paesaggio löschiano presenta altre interessanti peculiarità. In primo luogo in modo organico si registra un progressivo incremento del numero di funzioni e servizi man mano ci si allontana dalla località centrale. In questo senso, il modello introduce una nuova dimensione che regola in modo spontaneo la funzione della distanza tra agglomerazioni di servizi.

3.5 La regola rango dimensione

Solitamente, con il concetto "regola rango dimensione" si fa riferimento ad un modello geografico relativo alla popolazione degli insediamenti di un determinato territorio. Questo paradigma è conosciuto anche come legge di Zipf; da George Kingsley Zipf (1902-1950), linguista che identificò uno dei principi universali dell'organizzazione spontanea dei fenomeni naturali o culturali. In ogni caso, il primo ad avere formulato un paradigma geografico sulla relazione tra popolazione e rango è stato Auerbach nel suo breve saggio *Das Gesetz der Bevölkerungskoncentration* (1913), nel quale esponeva i principi relativi alla dimensione degli insediamenti in un determinato territorio.

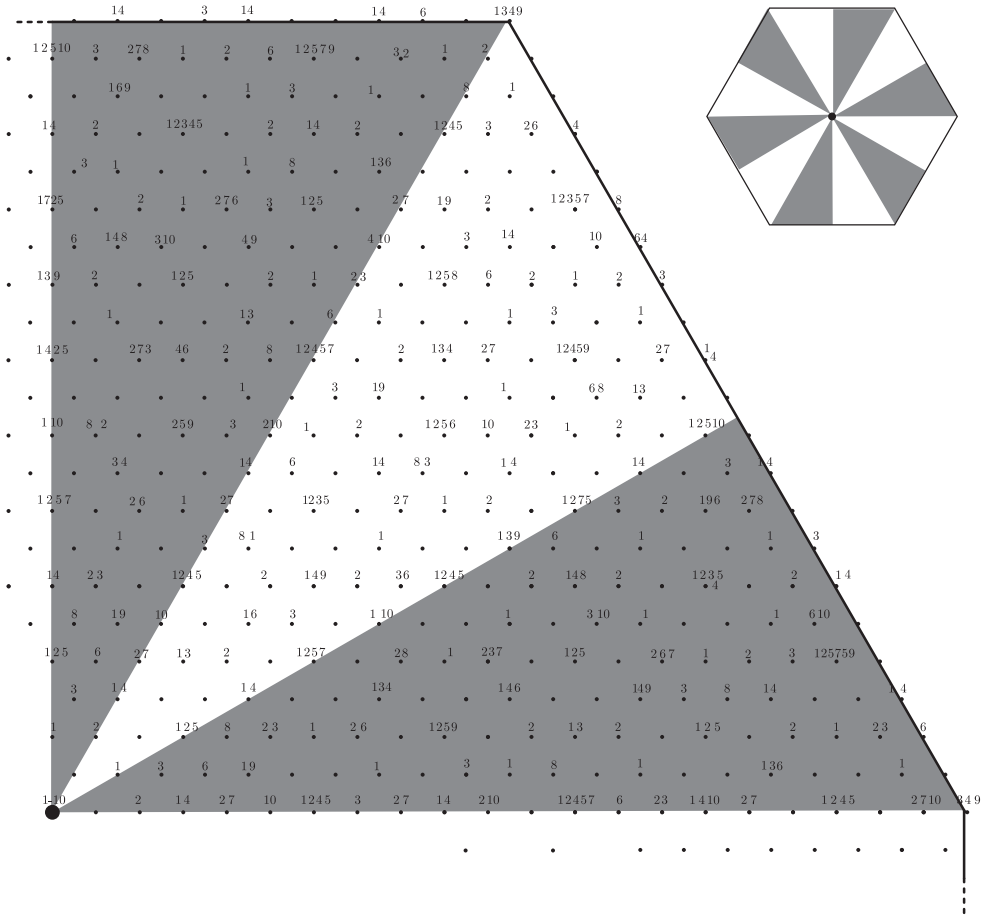


Fig. 3.24. Ricostruzione del paesaggio löschiano. Per ogni punto illustrato nel grafico sono indicati i numeri relativi alle prime 10 funzioni o ranghi secondo il modello di Lösch. Il luogo centrale presenta le funzioni da 1 a 10. Allontanandoci progressivamente da questo si può osservare un incremento delle funzioni o numero di servizi in certe località.

Le osservazioni di Auerbach corrispondevano ai *trend* del popolamento degli stanziamenti umani identificati e validati tempo dopo dal paradigma di Zipf. Comunemente Zipf viene indicato come padre di un modello che, pur non essendo nato nell'ambito della geografia, si sarebbe dimostrato di grande aiuto alla comprensione della logica dell'organizzazione della struttura gerarchica dei sistemi di stanziamento umano. La popolarità e diffusione del paradigma di Zipf fra i geografi sta nel fatto che con esso si identifica un principio in grado

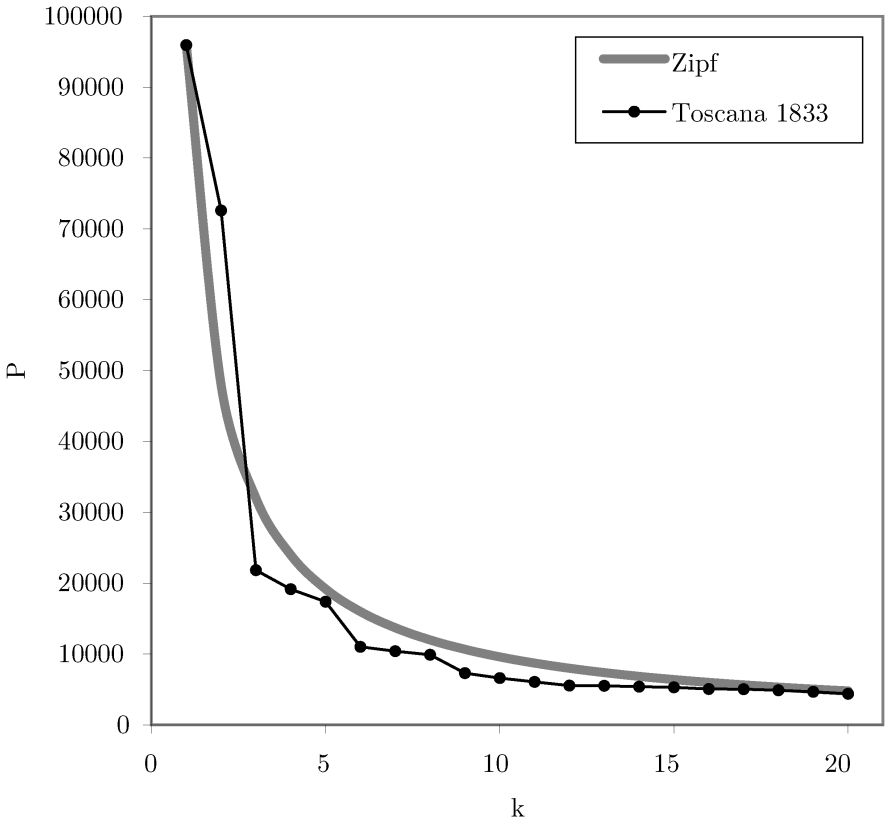


Fig. 3.25. Distribuzione della popolazione dei principali centri del Gran Ducato di Toscana e distribuzione di Zipf. Come si può notare dal grafico, l'andamento della popolazione reale nei diversi centri del territorio sotto esame, segue a grandi linee la distribuzione ideale della regola rango dimensione.

di descrivere in modo chiaro e formale l'organizzazione del popolamento di un determinato territorio[62].

Quello che maggiormente colpisce i ricercatori che si avvicinano per la prima volta a questo modello è la sua semplicità, e il poter verificare che, da uno schema così semplice, si riesca effettivamente a modellare l'andamento della popolazione dei diversi nuclei del territorio sotto esame. Oltre a ciò, va notato che l'applicabilità in ambito geografico del modello di Zipf è fondata sulle differenze e non sulle somiglianze a livello generale.

[62] Per una introduzione generale della regola rango dimensione consultare Haggett, Cliff, Frey 1977, pp. 110-112.

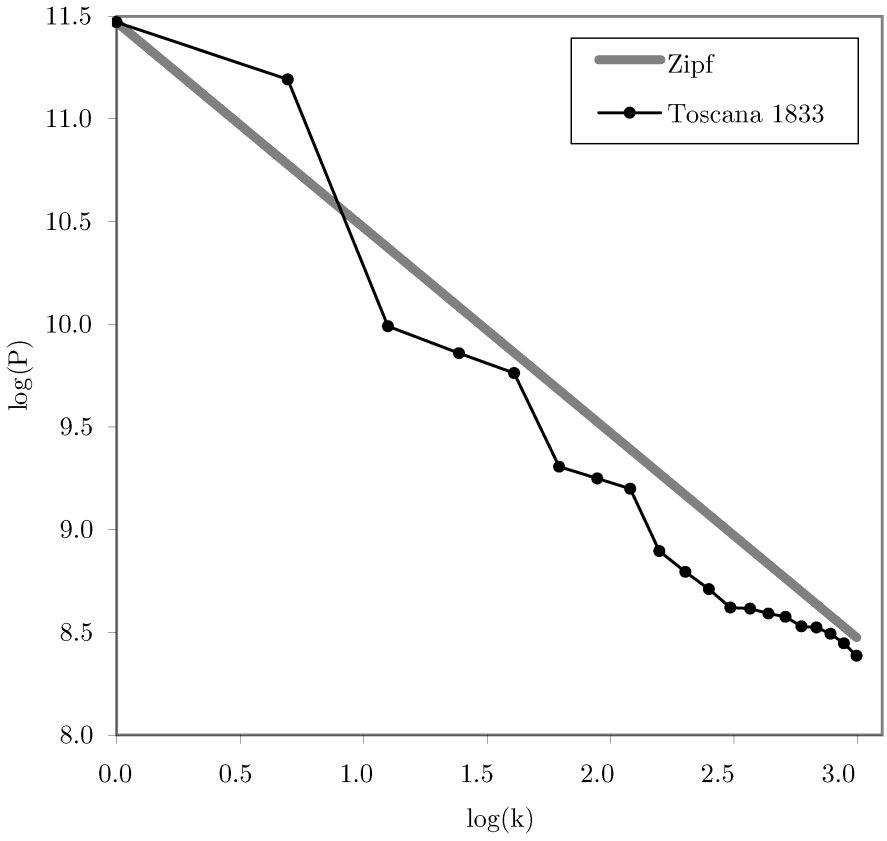


Fig. 3.26. In questo grafico sono stati riportati i logaritmi della posizione (k) e della popolazione (P). La strategia è importante perché, permettendo di rendere lineare il grafico, consente di enfatizzare le differenze tra il grafico idele (Zipf) e i valori osservati.

Alla base della relazione identificata da Auerbach e Zipf c'è un concetto elementare. La popolazione degli insediamenti in una certa regione si presenta decrescente a partire da quello più grande a quello più piccolo. Tale decremento (o sarebbe meglio dire, il ritmo di tale decremento) può essere modellato con la seguente funzione:

$$P_k = \frac{P_1}{k^{\Theta}} \tag{3.2}$$

Dove P_1 è la popolazione dell'insediamento più grande, P_k è quella del k insediamento (ad esempio P_3 è quella del terzo insediamento più grande), k è una posizione o indice di una sequenza ordinata di eventi e Θ è una costante

che determina il discostamento della curva ideale. La costante Θ è l'aspetto meno diffuso di questo approccio, ma è sicuramente uno dei più importanti in termini formali e di modellizzazione dei dati analizzati. Questo aspetto verrà discusso in seguito.

La distribuzione di Zipf, intesa come modello geografico, ha assunto il nome di regola rango dimensione, proprio perché introduce l'idea della dimensione attesa per un dato insediamento sulla base del suo rango (o posizione). Secondo questa regola, data una certa regione finita, e ordinando i singoli insediamenti al suo interno dal più grande al più piccolo (in termini di numero di abitanti), il secondo insediamento avrà teoricamente una popolazione ideale pari alla metà di quella del primo; il terzo una popolazione pari a un terzo di quella del primo e così di seguito. Ad esempio, immaginando che in un determinato territorio l'insediamento più grande abbia 25.000 abitanti, il modello di Zipf prevede che:

$$P_1 = \frac{25000}{1} = 25000$$

$$P_2 = \frac{25000}{2} = 12500$$

$$P_3 = \frac{25000}{3} = 8333$$

Questo principio, benché banale, descrive con un notevole grado di accuratezza la popolazione dei sistemi urbani di diverse aree del mondo e di diversi periodi storici. In questo senso, la regola rango dimensione, pur essendo un modello a una sola dimensione, va intesa come un paradigma generale della distribuzione spaziale del popolamento. Uno degli aspetti più interessanti di questo vero e proprio modello del popolamento è che esso si è dimostrato efficace nell'applicazione a ranghi insediativi inferiori, fornendo la possibilitàdi essere sfruttato in processi di analisi di aree rurali [63]. Non occorre qui insistere dunque sull'enorme importanza che esso potrebbe avere nel processo di comprensione dei paesaggi storici. Soprattutto se si considera che in una determinata regione, la dimensione di un dato insediamento – sempre secondo la logica del modello – è in qualche misura collegata a quella di tutti gli altri insediamenti[64]. In questo senso, il paradigma di Auerbach sembra confermare la teoria della natura sistemica dell'insediamento umano.

La regola regola rango dimensione è stata applicata in diversi contesti geografici; partendo dal livello globale, scendendo a quello continentale, nazionale

[63] Tim, Unwin 1981, pp. 352-355.

[64] Un resoconto di alcune implicazioni teoriche e speculative relative alla regola rango dimensione in ambito geografico si trova in Berry, Garrison 1958. Il modello presenta degli aspetti legati anche alla sua applicabilità nell'ambito dei frattali: Hsin-Ping 2004.

e regionale[65]. Qui si desidera porre l'accento alla sua applicabilità nei contesti regionali. E una dimostrazione dell'efficacia di tale modello si può cogliere dalla lettura della tabella 3.2, nella quale sono stati riportati il numero di abitanti dei primi 20 insediamenti del Granducato di Toscana agli inizi del XIX secolo (Repetti, 1833-1846). Come si può osservare dai dati della tabella, esiste una differenza (in alcuni casi marcata) tra l'andamento della popolazione reale (P) e i valori ipotizzati secondo la regola rango dimensione ($Zipf$). Dopo Firenze, i cui valori osservati e attesi sono uguali, dato che rappresenta il primo insediamento della serie, Livorno mostra una differenza di 24.621 unità. Progressivamente però tale differenza andrà diminuendo nella sequenza. In ogni caso, l'andamento dei valori delle due sequenze mostra come vi sia una correlazione positiva tra quelli osservati e quelli attesi. Nel primo grafico riportato nella figura 3.25 si può osservare l'andamento simile delle curve create con i dati provenienti dalla tabella. Si può notare infatti che le due curve, con la sola eccezione del secondo valore, appaiono caratterizzate da un andamento simile. Naturalmente quella grigia ha un andamento regolare dato che è il risultato della funzione matematica descritta precedentemente (3.2).

In ogni caso, lo scopo empirico più importante e utile ai fini dell'applicazione di questo modello sta proprio nell'identificare le differenze tra valori attesi e valori osservati e non tanto nelle similitudini, che in ogni caso tendono a persistere, come dimostra la figura 3.25. L'applicazione della regola rango dimensione su qualsiasi sistema di stanziamento reale comporta comunque l'identificazione di differenze tra valori osservati e ipotizzati.

Al fine di enfatizzare le differenze tra valori osservati e valori attesi si ricorre alla linearizzazione delle curve con l'aiuto delle scale logaritmiche. Nella figura 3.26 è riportato un secondo grafico per il quale sono stati calcolati logaritmi naturali (e) sia per le posizioni (k, sulle ascisse) che per il valore delle singole popolazioni (P_k, sulle ordinate)[66]. Di fatto, se non linearizzate, le due curve (valori attesi e valori osservati) tendono nella maggior parte dei casi a seguire un andamento molto simile, rendendo difficile qualsiasi identificazione prima, ed interpretazione poi, delle differenze.

Una volta linearizzata, l'inclinazione della curva, ovvero la forma della distribuzione, permette di identificare caratteristiche e attributi particolari (che qui definiremo socio-politici) del sistema di stanziamento. Nell'osservare il modo e la misura con cui le misurazioni si allontanano dai valori attesi si possono desumere informazioni sulle caratteristiche funzionali del sistema

[65] Per una panoramica sulla differenziazione delle curve rango dimensione sulla base dei contesti geografici e la loro evoluzione diacronica consultare Johnson 1980, pp. 235-236; Conway 1979. L'applicazione in contesti preistorici è stata affrontata da Pearson 1980.

[66] Il calcolo del logaritmo è stato realizzato in *Excel* con l'aiuto della funzione *log( )*.

Toponimo	***k***	***P***	***Zipf***
Fienze	1	95927	95927.0
Livorno	2	72584	47963.5
Lucca	3	21829	31975.7
Pisa	4	19144	23981.8
Siena	5	17378	19185.4
Prato	6	11014	15987.8
Arezzo	7	10402	13703.9
Pistoia	8	9892	11990.9
Fucecchio	9	7305	10658.6
Massa	10	6600	9592.7
Pescia	11	6068	8720.6
Empoli	12	5548	7993.9
Castiglion Fiorentino	13	5518	7379.0
Ponte Buggianese	14	5389	6851.9
Pontedera	15	5302	6395.1
Carrara	16	5063	5995.4
Cortona	17	5036	5642.8
Viareggio	18	4883	5329.3
Camaiore	19	4661	5048.8
Gangalandi (L. a Signa)	20	4388	4796.4

Tabella 3.2. Numero di abitanti (P) nei primi venti centri della Toscana nel 1833 secondo i dati riportati da E. Repetti nel *Dizionario Geografico Fisico Storico della Toscana*. Nella colonna *Zipf* vengono riportati i valori ideali secondo l'equazione 3.2.

gerarchico. In altre parole, notizie sui possibili effetti del sistema sociopolitico sul processo di articolazione della popolazione in quel determinato territorio.

In termini di lettura e analisi qualitativa dei dati, la regola rango dimensione è fondamentale dato che consente di stabilire un indice delle caratteristiche della distribuzione della popolazione all'interno della sequenza gerarchica. In altre parole, la regola rango dimensione permette di identificare le caratteristiche politiche che regolano la vita del territorio sotto esame. Dall'utilizzo della regola rango dimensione possono essere rilevate sostanzialmente tre condizioni o modelli interpretativi elementari:

- Modello binario (figura 3.27, b). Si verifica quando, una volta linearizzati, i valori osservati disegnano una curva al di sopra della curva ideale di Zipf. Tale caso è determinato da variazioni ridotte nella popolazione dei primi insediamenti della curva. Solitamente essi denotano uno stato

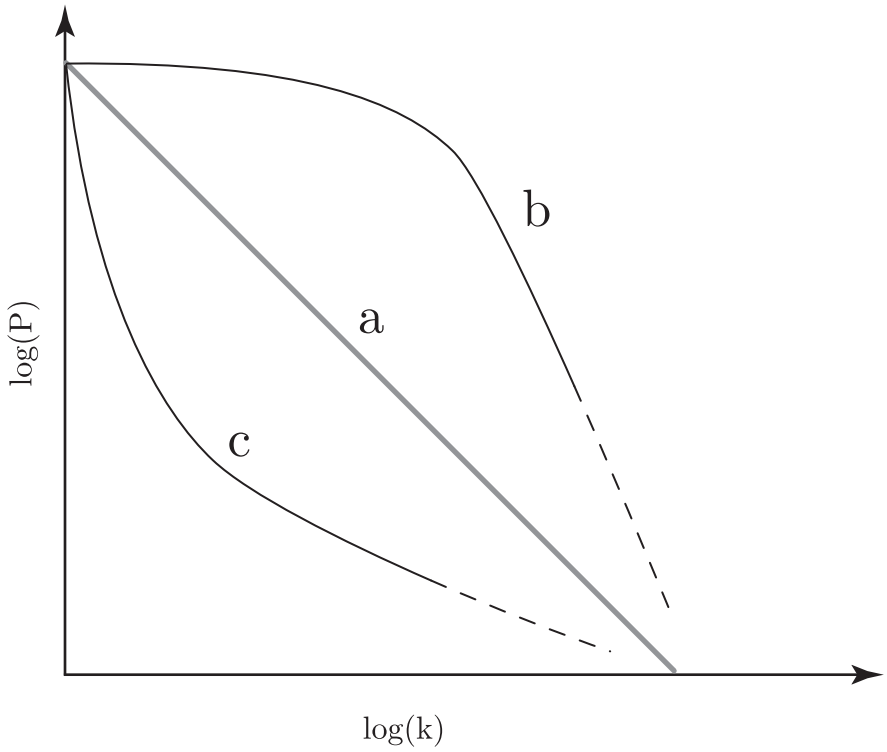


Fig. 3.27. Confrontando i valori osservati con quelli della curva di Zipf si verificano di norma tre casi: andamento al di sopra della curva ideale (a), noto come modello binario; corrispondenza (b); posizionamento al di sotto della curva ideale (c), solitamente definito modello delle città primate.

o forme amministrative caratterizzate da più capitali o, in ogni caso, da centri di controllo diversificati. Come per i sistemi federativi di cui l'esempio più emblematico sono le città etrusche, delle quali nessuna era la capitale o centro dominante. Il modello binario può essere definito "sistema anarchico". L'appellativo binario deriva dal fatto che non è un unico centro, bensì due o più, a dominare il territorio.

- Andamento naturale (figura 3.27, a). Si verifica quando la popolazione segue l'andamento ideale. In questo caso il sistema viene definito come andamento naturale o di mercato. In esso si presume che i fattori di libero mercato e dunque una economia di scambi naturali, con assenza di intervento statale, abbia lasciato crescere ogni singolo insediamento in modo naturale. In questo senso si può immaginare come traccia di processi politici di tipo democratico o, in ogni caso, di sistemi in cui le

leggi del libero mercato abbiano influenzato i processi di stanziamento.

- Modello delle città primate (figura 3.27, c). Si registra invece quando la curva dei valori della popolazione si colloca al di sotto di quella di Zipf. Solitamente questa condizione si verifica quando esiste un centro dominante che ha egemonizzato sia politicamente che socialmente l'intero sistema sottostante. In altri termini, esiste una differenza molto marcata tra la prima città e il resto della sequenza ordinata. In questo caso si è in presenza di un sistema retto da un grande nucleo principale, con una popolazione fuori scala rispetto al resto del sistema di stanziamento; caso tipico dei sistemi politici di tipo imperiale.

Riallacciandoci con quanto detto prima, la funzione ha una costante Θ che regola la forma o inclinazione della curva ideale linearizzata. Essa viene impiegata come indice del grado di inclinazione e classificazione della curva osservata. Se la curva dei valori osservati ha un andamento molto simile a quella dei valori attesi allora $\Theta \approx 1$. Nei casi in cui si registrino delle differenze marcate tra valori osservati ed attesi, si possono avere due casi diversi: di fronte ad un sistema binario $\Theta > 1$, mentre, per uno primate, $\Theta < 1$.

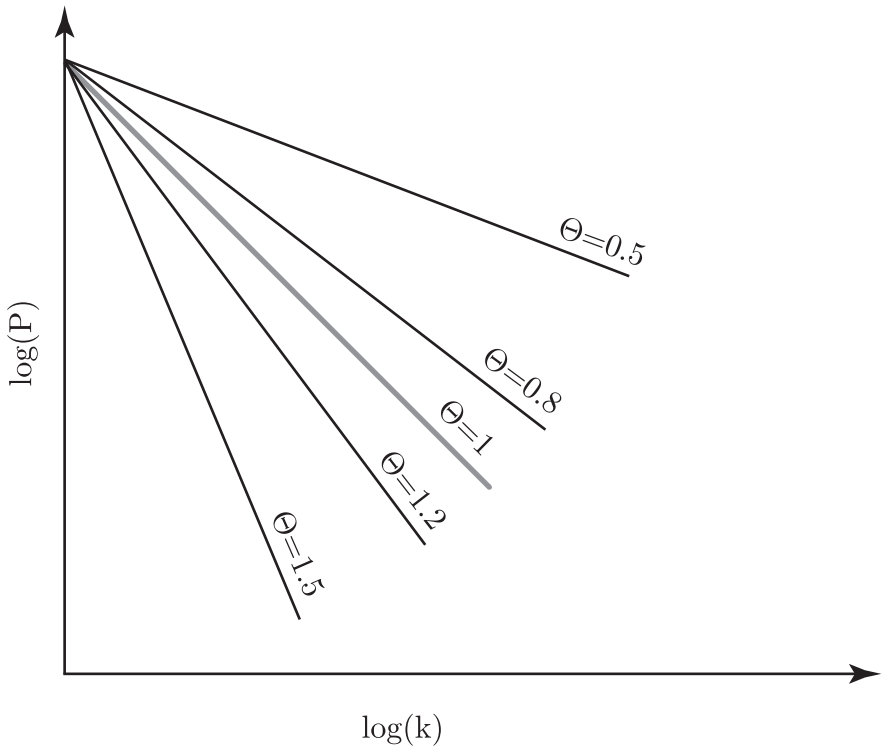


Fig. 3.28. Nel modello di Zipf, il valore di Θ può essere utilizzato per modellare l'inclinazione della curva ideale. Ai fini pratici, il valore di *Theta* può essere effettivamente utilizzato per identificare le differenze tra valori osservati e attesi, cioè quelli con $\Theta = 1$.

4
Introduzione alla statistica descrittiva in geografia

4.1 Le misurazioni

Il dizionario Treccani definisce la "misurazione" come il processo di «determinare la misura di una grandezza; eseguire una misurazione; prendere (o trovare, stabilire, calcolare) le misure di qualche cosa». Cercando di espandere tale definizione si può aggiungere che il processo di misurazione ha come scopo quello di compiere una o più osservazioni all'interno di un "quadro di riferimento". Nel caso della misurazione scientifica – qui intesa come quella formale e critica – il quadro di riferimento deve naturalmente essere condiviso dai membri che compongono la comunità all'interno della quale i dati raccolti devono essere distribuiti e fruiti[1].

Se la misurazione può essere associata al processo d'acquisizione di informazioni, le "misure" rappresentano la base elementare del linguaggio con il quale più soggetti appartenenti ad un'unica comunità dialogano o possono in ogni caso legittimare qualità e significatività dei dati scambiati[2]. Appare evidente che solo se i membri della comunità condividono questo quadro di riferimento sarà possibile il dialogo e lo scambio di idee. Questo equivale a dire: *se hai usato questi parametri di misurazione, (comuni anche al mio modo di procedere), io posso, in linea di principio, utilizzare e valutare i dati che tu hai*

[1] Ballard 1949 p. 134; Dingle 1950, p. 7.

[2] Byerly, Lazara 1973, p. 10: «Some aspects of measurement are clearly conventional in the strong sense of resting on essentially arbitrary agreements. Thus to ask which of the scales of temperature Centigrade of Fahrenheit is the "true" scale would be like asking which are the true words for the first three integers: *one*, *two*, *three* or *eins*, *zwei*, *drei*. [...] Conventions consist in agreements among people which reflect decision that "could be otherwise". The question is: on pain of what can a convention be maintained?»

registrato. In questo contesto, misura e misurazione acquistano due significati molto diversi [3].

La geografia quantitativa e l'analisi spaziale non sono altro che un corpo di procedure di misurazione all'interno di un contesto di tipo spaziale[4]. In altri termini la traduzione dell'osservazione diretta della realtà geografica in rilevamenti formali. Quello che sorprende maggiormente, come si cercherà di illustrare di seguito, è come la quasi totalità delle osservazioni relative al rapporto tra specie umana e ambiente costituisca già o possa essere facilmente convertita in misure formali.

Spesso si pensa che le osservazioni critiche sull'uomo non siano traducibili in un linguaggio formale. Questo perché la sfera antropologica appare particolarmente arbitraria e legata a schemi erratici. Occorre considerare invece come i fenomeni sociali e antropologici non sono meno caotici di gran parte del mondo fisico e naturale.

Uno dei vantaggi principali dei quadri di riferimento scientifici è la loro semplicità. Una cornice di riferimento dove condurre processi di misurazione si può immaginare composta principalmente da due elementi: il "campione" e la "scala di misurazione"[5]. Questo vuol dire che la significatività – ed anche il senso – di un insieme di misurazioni compiute da altri interlocutori può in prima istanza essere stabilita in base alle scale di misurazione e al processo di campionamento adottate nella fase della loro raccolta.

È facile intuire come sia impossibile condurre processi di quantificazione geografica al di fuori della logica e della teoria della "misurazione scientifica". Questo è il punto centrale e la sintesi dell'idea che con questo testo si cerca di illustrare. La forza della misurazione dello spazio sociale non sta in una presunta sistemazione di tipo scientifico quanto nella possibilità di condividere in modo preciso e puntuale, con una comunità potenzialmente vasta, non solo il corpo delle misurazioni raccolte durante le proprie ricerche, ma anche e perfino l'evoluzione del proprio pensiero e le proprie interpretazioni[6].

Misurare, in ultima istanza, significa comunicare. Rendere esplicite e chiare le proprie osservazioni sul mondo e facilitare agli altri il processo di comprensione delle nostre interpretazioni della realtà. «I metodi quantitativi costituiscono solo uno schema di riferimento per la *concettualizzazione di un problema e per*

[3] Dingle 1950, pp. 12-13: «I think the reason why we make some measurements and not others can be understood only in terms of the ultimate object of science, which is to find relations between the elements of our experience. If we passively accept experience as it comes the problem is far too difficult, so we establish artificial conditions that yield results more easily related with one another than natural occurrences are.»

[4] Matthews 1985, pp. 15-16; Ripley 1981, pp. 1-9; Unwin 1986.

[5] Da non confondere con le scale geografiche.

[6] Coppock, Johnson 1962, pp. 130-131.

la sua formalizzazione teorica. Gli strumenti matematici intervengono soltanto come un linguaggio più comodo, più efficace perchè più preciso rispetto a quello verbale.»[7].

4.1.1 La misurazione spaziale

Il caso dell'analisi spaziale presenta un'ulteriore particolarità. Nella quasi totalità dei casi, le misurazioni all'interno di un processo d'analisi spaziale avvengono attraverso la mediazione di una carta geografica; oggigiorno quasi sempre di tipo digitale[8]. In molti casi il ricercatore raccoglie le osservazioni ed i dati per le proprie analisi non direttamente dalla realtà, ma attraverso l'utilizzo di supporti cartografici di diversa natura. Questo significa che il ricercatore che intenda avviare processi d'indagine quantitativa deve confrontarsi comunque con livelli d'interpretazione della realtà compiuti nella maggior parte da altri individui. La carta geografica non equivale alla realtà. Essa costituisce più semplicemente una interpretazione realizzata da altri attraverso processi cognitivi di tipo selettivo e discriminanti.

Anche se in linea di principio si può accettare che la rappresentazione cartografica abbia raggiunto un elevato grado di precisione, bisogna sempre ricordare che, in ogni caso, lo studio delle forme di distribuzione spaziale costituisce sempre un processo indiretto di osservazione della realtà mediato dal supporto cartografico[9].

4.2 Le scale di misurazione

L'attività di misurazione ha come scopo quello di mettere in relazione un'osservazione con una particolare scala di riferimento. Questo equivale ad assegnare un certo valore o misura agli oggetti studiati[10]. Per misura intendiamo dunque la traduzione in linguaggi formali o quantitativi di attributi specifici comuni ad un insieme di oggetti. Per questo motivo, le scale di misurazione

[7] Merlin 1983, p. 210.

[8] Per un esempio chiaro dell'evoluzione dell'analisi spaziale da una metodologia classica a un'era post-GIS confrontare Unwin 1986 e O'Sullivan, Unwin 2003. È molto interessante osservare l'introduzione del concetto di *Geographic Information Analysis*; ancora più rilevante se si prendono in considerazione le osservazioni di Wright, Goodchilde, Proctor 1997 relative alla natura o al ruolo dei GIS.

[9] Vedi concetto di "modello": capitolo 2. L'analisi spaziale non si limita esclusivamene alla misurazione degli aspetti distributivi dei fenomeni. Certo è che il processo di analisi spaziale deve se non altro contare sul processo di raccolta, ordinamento e analisi di variabili riferite alla dimensione spaziale. È convenzione quasi universale associare tale rappresentazione delle coordinate con le variabili x e y.

[10] Galton 2000, pp. 6-8.

costituiscono la base fondamentale dell'intero processo di quantificazione. Se l'approccio quantitativo rappresenta un particolare tipo di linguaggio, le scale a loro volta possono essere immaginate come i loro possibili abecedari. E conoscere e comprenderne le differenze costituisce un fattore essenziale.

Il processo di misurazione può essere immaginato come la traduzione di un "attributo" specifico in un valore della scala di riferimento[11]. Non si misura la persona, bensì il suo peso o la sua età. Solitamente all'attributo misurato viene assegnato il nome di "variabile". Così attributi come peso, colore o dimensione, divengono, nel linguaggio corrente della ricerca scientifica, le variabili "peso", "colore" e "dimensione".

Una delle funzioni delle scale è quella di poter compiere operazioni di confronto tra più misurazioni. Ad esempio, se si misura l'estensione di due insediamenti, si può stabilire in modo preciso il rapporto tra i due; in altri termini, è possibile formulare affermazioni del tipo *l'insediamento A è 1.8 volte più grande del B*. Misurare significa infatti anche potere comparare in modo preciso due osservazioni[12].

Sempre in base a quanto fin qui illustrato, si può affermare come l'unico elemento distintivo tra un'osservazione e una misurazione sia che quest'ultima è sempre e in ogni caso riportata ad una scala di riferimento nota sia a colui che compie la misurazione, sia ai suoi destinatari. Vale a dire che le informazioni vengono codificate e decodificate secondo quel linguaggio comune di cui si parlava sopra.

Solitamente le scale di misurazione possono essere classificate in quattro categorie principali: nominali, ordinali, a intervalli e a rapporti. Saranno gli attributi stessi a suggerire con quale di queste scale le misurazioni dovrebbero essere compiute[13]. In altre parole, esiste sempre una scala ideale per ogni tipo di variabile[14]. Il lettore, peraltro, non si deve fare ingannare dal presunto grado di precisione delle scale. Ad esempio, a prima vista potrebbe sembrare che la scala a rapporti sia più precisa di quella nominale; questo però non è assolutamente vero. Certo risulterebbe molto difficile calcolare le quote sul livello del mare con una scala nominale; ma è anche vero che lo sarebbe altrettanto misurare il tipo d'uso del suolo con una scala a rapporti al posto di

[11] O'Sullivan, Unwin 2003, p. 11: «When information is collected, measurement is the process of assigning a class or value to an observed phenomenon according to some set of rules. It is not always made clear that this definition does not restrict us to assignments involving numbers. The definition also includes the classification of phenomena into types, or their ranking relative to one another on an assumed scale.»

[12] Mitchell 1994, pp. 389-391.

[13] Per una miglior comprensione dell'uso delle quattro scale di misurazione in ambito cartografico e geografico si rimanda a Coulson 1978, pp. 48-64.

[14] Per il concetto di variabile vedasi *infra*.

quella nominale. Di fatto, gran parte delle misurazioni geografiche si basano sulla scala nominale e questo non implica un minor grado di precisione.

La comprensione della natura e delle caratteristiche delle scale di misurazione diviene fondamentale non solo in termini di misurazione all'interno di processi di ricerca condotti nell'ambito della geografia quantitativa o dell'analisi spaziale, ma soprattutto nella produzione cartografica. Di frequente si trovano carte costruite e realizzate sulla base di rappresentazioni simboliche codificate con la scala sbagliata, oppure con un uso non appropriato. Se ci fosse bisogno di aggiungere ulteriori motivazioni al perché della geografia quantitativa, sarebbe sufficiente considerare che la produzione cartografica, al di là di quella squisitamente topografica, richiede sempre l'utilizzo di un linguaggio formale per la codifica simbolica. Dominare il linguaggio formale, significa essere pienamente in grado di convogliare nel migliore modo possibile il proprio messaggio attraverso la carta geografica[15].

4.2.1 La scala nominale

La scala nominale rappresenta lo schema più semplice per definire e valutare le osservazioni[16]. Questo tipo di scala ha come scopo quello di misurare le differenze di categoria o "classe" e non di grado o peso. In altre parole, nella scala nominale, ogni oggetto viene classificato in base alla sua "appartenenza" ad una categoria. Ad esempio, classificare un determinato gruppo in "maschi" e "femmine".

Come è stato suggerito sopra, la sua semplicità non deve essere confusa con una forma di scarsa precisione o efficacia nell'organizzazione dell'informazione. Semplicemente, la scala nominale non prevede relazioni di tipo quantitativo né qualitativo fra le classi enunciate. I valori osservati costituiscono semplicemente una categoria, gruppo o classe a sé stante. Proprio per tale motivo, variabili espresse in scala nominale contengono sostanzialmente valori che esprimono il nome della categoria d'appartenenza. Nella geografia umana e fisica le scale nominali sono fondamentali; ed è proprio per questo, ad esempio, che in cartografia se ne fa largo uso.

Nella geografia, i valori di una scala nominale possono essere le province di una determinata regione; ad esempio la Lombardia: *Varese, Como, Lecco,*

[15] Haggett, Cliff, Frey 1977, p. 291: «The types of maps that geographers can draw are directly controlled by the level of the measurements that they can collect. The notion of "levels of measurement" is one that has fundamental applications in statistics, but its importance in mapping has been generally overlooked...». Su questo tema fondamentale si consiglia di consultare anche Dent 1972, p 93. Vedi *infra* al paragrafo 7.1.

[16] Unwin 1986, p. 65; Silk 1979, p. 6; O'Sullivan, Unwin 2003, p. 12; Matthews 1985, p. 16.

Fig. 4.1. Le scale di misurazione sono organizzate in nominale, ordinale, ad intervalli e a rapporti.

Sondrio, Milano, Bergamo, Brescia, Pavia, Lodi, Cremona e Mantova. Una volta definita questa scala, si può procedere semplicemente a misurare il valore dei centri abitati in base alla loro localizzazione entro i confini di una o dell'altra provincia. Ad esempio, utilizzando questa scala, come risultato della misurazione si ottiene:

$$\begin{aligned} Voghera &\Rightarrow Pavia \\ Saronno &\Rightarrow Varese \\ Monza &\Rightarrow Milano \\ Gallarate &\Rightarrow Varese \\ Varese &\Rightarrow Varese \end{aligned}$$

All'interno di questa scala, perciò, non vi è nessun riferimento alle qualità, pesi o attributi delle singole provincie. Il fatto di avere registrato per Voghera il valore di "Pavia" non significa registrare o attribuire un valore alle qualità intrinseche di questa città; ci si limita esclusivamente a considerare tale centro come appartenente alla classe di città e centri appartenenti alla provincia di Pavia. L'associazione di elementi cartografici ai valori di questa scala non ha altro scopo che quello di indicare l'appartenenza di questo o quell'oggetto ad una determinata provincia. La scala nominale non esprime di per sé differenze né di grado, né di gerarchia o peso. Essa non suggerisce nemmeno l'estensione dell'area del territorio, numero di abitanti o grado di sviluppo economico. Essa indica semplicemente le diverse categorie, tipi o classi di province che compongono la regione Lombardia. Nella figura 4.2 i nomi sono esposti in disordine, ad indicare che non esiste proprio nessuna forma o possibilità di assegnare un peso o posizione a tali misurazioni.

In geografia, altre scale di tipo nominale possono essere i comuni che compongono una provincia, i tipi geologici, litologici, l'uso de suolo, ecc.

Una delle caratteristiche dei valori all'interno di una scala nominale è che essi possono essere trasformati liberamente senza che per questo si perdano l'essenza stessa e la funzionalità della scala. Di fatto, i valori all'interno di una scala nominale non hanno altro significato che quello di etichetta. Così ad esempio *Bergamo, Brescia, Como...* possono diventare *Classe F*, *Classe G*, *Classe M...* senza per questo rendere meno effettiva la funzionalità della scala. Naturalmente spetta a chi compie l'operazione di definire la convenzione con la quale tale scala può essere decodificata.

La scala nominale può apparire molto limitata ai fini di diversi passaggi di quantificazione statistica, ma, nonostante ciò, molti dei processi di analisi spaziale possono sfruttarla. Ciò è particolarmente vero per quanto riguarda analisi di frequenza e studi di densità. Di fatto, la scala nominale esprime,

da un punto di vista spaziale, il principio di localizzazione, appartenenza, o presenza all'interno di un confine politico o amministrativo.

4.2.2 La scala ordinale

Il secondo tipo di scala è quello ordinale. Come indica il suo stesso nome, la scala ordinale, in relazione a quella nominale, introduce nel processo di misurazione il concetto di ordine[17].

Essa è composta da classi che possono essere organizzate in una sequenza che indica in qualche misura categorie ordinate anche in base ad una sequenza gerarchica. Questo ordine non esprime però i rapporti specifici che intercorrono tra le diverse classi. Un esempio di scala ordinale è: *primo, secondo, terzo, quarto ...* ecc. Oppure *ottimo, buono, mediocre, ...* Tali scale esprimono l'accrescere di un certo valore o attributo specifico; ma non rendono esplicito di quanto.

Per questo motivo, nel caso tale scala venga utilizzata, i dati non potranno essere distanziati, ovvero "pesati". Ad esempio, appartenere alla categoria *ottimo* non significa essere due volte *buono* o tre volte *mediocre*. Anche nel caso specifico tutto questo fosse vero, la scala ordinale non esprimerebbe comunque tali rapporti interni; non è il suo scopo.

Le scale ordinali godono della proprietà transitiva all'interno delle proprie relazioni. Così, se la *Classe 1* precede la *Classe 2* e la *Classe 2* precede la *Classe 3*, allora anche la *Classe 1* precederà la *Classe 3*.

Nella geografia le scale ordinali sono abbastanza comuni. Una scala ordinale potrebbe essere: *pianura, collina, montagna*; oppure, a livello di gerarchia politica: *comune, provincia, regione*. Una scala come *alto, medio e basso* viene usata spesso nella geografia umana in riferimento ad attributi specifici di un gruppo sociale, come ad esempio benessere economico, criminalità, salute, oppure per indicare aspetti come il grado d'inquinamento dell'aria.

Le scale di tipo ordinale hanno però un impiego limitato nel processo di analisi dei dati spaziali. Di fatto, la scala ordinale esprime in qualche misura una visione particolare della realtà. La soglia del livello di benessere economico tra *ricchi* e *classe media* può cambiare a seconda del gruppo di ricerca, il luogo, e l'epoca storica sulla quale si compie la misurazione. Questa condizione fa sì che l'impiego all'interno di processi di misurazione possa essere condizionato e limitato dalle posizioni relative implicite in essa. Senza voler conferir loro una valenza particolarmente negativa, si può affermare che nel processo dell'analisi spaziale dei dati, l'utilizzo di misurazioni compiute su scale ordinali è sempre un problema. Quella ordinale spesso possiede una natura soggettiva.

[17] Haggett, Cliff, Frey 1977, p. 292; Unwin 1986, p. 66; Matthews 1985, p. 16.

Nonostante questo, il concetto di scala ordinale è molto diffuso nella fase di esposizione dei risultati analitici; fase nella quale, una volta conclusa quella di quantificazione, sono richiesti anche strumenti in grado di esprimere in modo sintetico, semplice e comprensibile i risultati misurati.

In termini puramente cartografici, è chiaro dunque che la scala ordinale, in un certo modo, richieda un processo di rielaborazione e interpretazione dei dati. La sua funzione resta quasi esclusivamente relegata all'ambito della presentazione o illustrazione dei dati. In termini analitici costituisce forse una delle scale meno diffuse, dato che sottintende un livello di arbitrarietà difficilmente gestibile, sia da parte di chi compie le analisi che da parte di chi successivamente ne fa uso.

4.2.3 Scala ad intervalli

La scala ad intervalli costituisce un sistema di misurazione in cui è possibile stabilire le differenze in termini numerici tra due rilevamenti[18]. Vale a dire che i valori non solo sono ordinati come nel caso precedente, ma in questa scala le misure sono distanziate ad "intervalli" regolari identici tra di loro. Grazie alla definizione di questi parametri è possibile compiere confronti di peso tra due o più misure. Così l'intervallo che separa 10°C da 20°C è lo stesso che separa 30°C da 40°C.

Le scale ad intervalli non dispongono però di uno zero naturale o assoluto. Per questo motivo non è possibile definire i rapporti fra le misurazioni. Tali scale cioè non permettono di definire la proporzione o rapporto tra una misura e l'altra. Un esempio molto comune di questa tipologia sono le scale termiche. Nel caso della scala Celsius, pur essendoci un valore 0, questo non indica affatto l'assenza totale di temperatura: cioè 0°C indica comunque un certo valore termico che convenzionalmente è stato posto nel punto di congelamento dell'acqua al livello del mare. Così -5°C è una temperatura minore di 0°C, −10°C minore di -5°C, e così di seguito. Questi valori indicano in ogni caso una certa quantità di energia termica. Il fatto che non esista uno 0 assoluto fa sì che non sia possibile affermare che 20°C sia, in valori termici, il doppio di 10°C. L'assenza di uno zero assoluto impedisce appunto di stabilire il rapporto fra misurazioni compiute con questo tipo di scala.

Un altro esempio di scala ad intervalli sono i giorni, le settimane, i mesi, gli anni, misurati all'interno di un calendario. Infatti, benché sia ben evidente che 20 anni sono il doppio di 10, non è possibile affermare che il 200 d.C. sia la metà del tempo se confrontato al 400 d.C. Potrebbe esserlo se confrontato alla

[18] Unwin 1986, p. 66; Silk 1979, pp. 6-7; O'Sullivan, Unwin 2003, p. 13; Matthews 1985, p. 16.

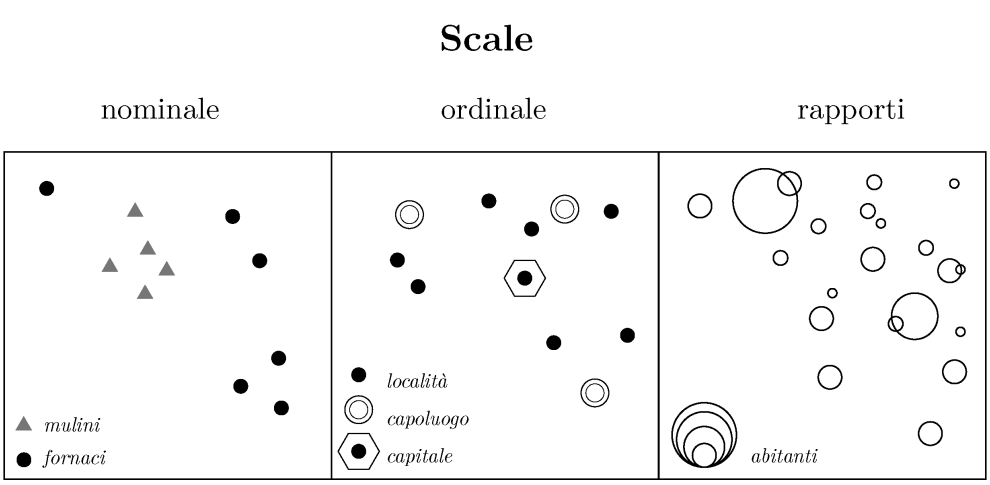


Fig. 4.2. Esempi a livello cartografico di scale di misurazione nominale, ordinale e a rapporti. Sul ruolo delle scale di misurazione nella rappresentazione dei fenomeni consultare Haggett, Cliff, Frey 1977, pp. 291-297 (figura 9.1 p. 292); Unwin 1986, p. 62-77 (in particolare figura 2.2, p. 76).

nascita di Cristo, che è una data puramente arbitraria. Ma tale affermazione non avrebbe alcun senso se riportata in termini di origine dell'uomo o nascita del sistema solare.

Ciononostante, con questo tipo di scala, oltre a definire l'ordine delle classi (come nel caso di quella ordinale), grazie all'introduzione di valori numerici graduati su intervalli identici, è possibile definire le distanze che intercorrono tra esse. Questo fattore è alquanto importante. Infatti, anche se non ci è consentito affermare che, in termini calorici, 20°C sono il doppio rispetto a 10°C, è anche vero che, se la temperatura in una stanza passa da 20°C a 25°C, è possibile a quel punto affermare che c'è stato un aumento di 5°C. In altri termini, pur non permettendo di stabilire rapporti o proporzioni, le scale ad intervalli consentono di definire i valori o attributi espressi in termini di posizione all'interno della scala.

4.2.4 Scala a rapporti

Anche nel caso della scala a rapporti, le classi sono distanziate con intervalli identici ma, a differenza del tipo di scala appena illustrata, questa possiede uno

zero assoluto[19]. Ad esempio, in una scala metrica, 0 metri indicano l'assenza totale di distanza. Questo valore nullo (o zero assoluto, appunto) rappresenta il punto di riferimento che consente di realizzare confronti e stabilire "rapporti" fra due o più misure. Grazie allo zero assoluto è possibile affermare dunque che 100 metri è, effettivamente, il doppio di 50 metri.

In geografia, scale a rapporti possono essere appunto la quota sul livello del mare, o la popolazione dei centri abitati, l'area dei comuni o la distanza in linea d'aria da un punto all'altro. Insieme a quella nominale costituisce la scala più diffusa in ambito cartografico.

4.3 Le variabili

All'interno di un quadro di riferimento scientifico, la raccolta di misurazioni si traduce nella definizione e nell'acquisizione di una "variabile". Le variabili possono essere definite una caratteristica o attributo che muta a seconda dell'elemento o oggetto sul quale viene compiuta la misurazione[20]. Ad esempio, la popolazione o il numero di abitanti per insediamento è una variabile: nel processo di censimento viene misurato tale valore per ogni singolo insediamento di un certo territorio.

L'intero quadro delle attività tipiche dell'applicazione dei metodi quantitativi in geografia si può concepire come un processo di raccolta, elaborazione e analisi di variabili. Nella geografia, infatti, la quasi totalità delle osservazioni o misurazioni è traducibile in variabili: distanza dai corsi d'acqua, pendenza dei suoli, distanza dalla città più vicina, popolazione, densità media di abitanti per chilometro quadro, indice di dispersione dell'insediamento, livello di formazione scolastica, indice di produttività delle aziende agricole, indice di fertilià dei suoli ecc. Così, molti aspetti di una regione o di un paesaggio divengono fattori quantificabili sotto forma di variabili. Vi sono molti più attributi fisici e sociali codificabili e gestibili sotto forma di variabile di quanto non ci si possa immaginare. È proprio per questo motivo che la corretta applicazione dei metodi, non solo di quantificazione ma anche di "rappresentazione" dai dati osservati, ha tanta importanza.

[19] Unwin 1986, pp. 66-68; Silk 1979, p. 7; O'Sullivan, Unwin 2003, pp. 13-14; Matthews 1985, p. 16.

[20] Per una panoramica introduttiva al concetto delle variabili consultare Bohrnstedt, Knoke 1994, pp. 25-27.

4.3.1 Variabili categoriali o mutabili

Le variabili non sono esclusivamente numeriche. Anche attributi misurati con l'aiuto di scale nominali possono essere registrati come "variabili categoriali" o "mutabili", come ad esempio uso del suolo, geologia o litologia. Questa costituisce una prima distinzione. Infatti, valori raccolti con scale nominali e ordinali sono definiti "variabili qualitative", mentre, nel caso di osservazioni compiute con l'aiuto di scale ad intervalli e a rapporti, si definiscono "variabili quantitative". Questa distinzione e soprattutto il concetto di qualitativo non devono trarre in inganno il lettore. Come verrà spiegato di seguito, entrambi i tipi di variabili possono essere impiegati in modo efficace all'interno di processi di analisi del territorio.

4.3.2 Variabili numeriche e quantitative

Le "variabili quantitative" vanno distinte a loro volta in due diverse tipologie: "variabili continue" e "variabili discrete". Tale distinzione è per alcuni versi forzata ma necessaria nel processo di comprensione delle modalità e della logica di certi processi di quantificazione. In generale si può affermare che una variabile continua può essere concepita come relativa a valori che cambiano, come indica il nome stesso, in modo continuo; in altri termini, per esprimerla, bisogna fare ricorso ai numeri reali ($\mathbb{R}$). Un esempio a livello geografico è la quota sul livello del mare. Infatti i possibili valori variano in modo continuo dato che un certo punto sul territorio può essere a 560 metri s.l.m., ma potrebbe essere anche a 562,34 metri s.l.m. Una variabile discreta, al contrario, va intesa come attributo codificabile sempre sotto scala numerica, ma che può cambiare per quantità fisse. Solitamente queste variabili vengono registrate grazie all'uso dell'insieme dei numeri naturali ($\mathbb{N}$). Nel campo della geografia quantitativa queste sono variabili molto diffuse. Ad esempio, dato un certo numero di villaggi in un territorio, per ognuno di questi si potrebbe registrare il numero di centri vicini entro un raggio di 5 chilometri; potrebbero essere 1, 2, 3 ... ma certamente non 2,36.

4.3.3 Notazione delle variabili

Solitamente alle variabili viene assegnato un nome o "simbolo". Si tratta di lettere del tipo A, B, C, X alle quali si attribuiscono i valori relativi ad un attributo specifico rilevati per gli elementi di un campione sotto esame. Così, ad esempio, quando ci si riferisce al numero di abitanti di un certo territorio, si può fare uso del simbolo P. Tutti i valori della colonna "popolazione (1951)"

nella tabella 4.1 possono essere rappresentati da questo simbolo:

$$7053, 9734, 4444, 5263, \cdots$$

Questo gruppo di valori corrisponde alla variabile P, in considerazione del fatto che essa assume, come detto sopra, una molteplice (n) quantità di valori. Perciò si può affermare che la variabile P equivale a

$$P_1 = 7053, P_2 = 9734, P_3 = 4444, P_4 = 5263, \cdots P_n = 3519$$

Spesso, all'interno di una formula, grafico o tabella, si usa la notazione P_i per indicare tale sequenza. La lettera i assume semplicemente il significato di di "indice" della variabile. Quest'ultimo rappresenta la sequenza di valori 1, 2, 3, 4, ... n. Perciò, ad esempio, con la notazione formale si vuole semplicemente esprimere

$$\text{popolazione comuni provincia di Siena} = P_i$$

In diversi settori della matematica i simboli i, j o k vengono utilizzati esclusivamente per indicare tali indici.

4.4 Il campionamento

L'obbiettivo di qualsiasi processo d'indagine scientifica è la produzione o l'acquisizione di nuova conoscenza su un determinato aspetto della realtà. Nella maggior parte dei casi, questi processi sono riferiti a oggetti e fenomeni distribuiti in termini di estensione e densità nel tempo e nello spazio. Lo studio di qualsiasi oggetto (come ad esempio i corvi o i cigni) presuppone l'indagine non tanto di un unico esemplare quanto piuttosto dell'intera categoria. Conosciamo qualcosa, ad esempio, sui corvi nella misura in cui conosciamo il fenomeno nella sua complessità.

È facilmente intuibile però capire come sia impossibile giungere alla formulazione di affermazioni di tipo assoluto su tali uccelli; questo sarebbe possibile solo dopo l'ispezione di ogni singolo esemplare[21]. Così, per giungere all'affermazione *tutti i corvi sono neri*, è necessario compiere una rilevamento su di un gruppo di esemplari sufficientemente ampio per arrivare alla formulazione di un concetto sul colore di questi volatili.

Il campionamento è il processo di selezione degli esemplari da osservare o sottoporre al processo di misurazione. Scopo del campionamento è normalmente quello di ottenere un "campione rappresentativo" di un'intera popolazione[22]. Nel fare ciò ci si deve assicurare che il campione sia di massima

[21] Vedi *supra* (capitolo 2) il paradosso di Hempel.
[22] Matthews 1985. p. 17.

rappresentatività; nel senso che esso deve in qualche misura essere imparziale nella rappresentazione della popolazione stessa. Il campione, fra l'altro, deve tenere conto anche delle caratteristiche distributive in termini spaziali del fenomeno studiato, in modo tale appunto da far sì che il campione scelto sia sufficientemente significativo. Ad esempio, sarà necessario osservare un numero consistente di corvi e comunque rappresentativi di tutti i continenti per giungere ad una asserzione assoluta come quella esposta sopra. Diversamente, si potranno formulare solamente affermazioni relative, del tipo *i corvi di questo territorio sono neri.*

Il processo di campionamento viene applicato sulla "popolazione" o "universo". Questi due termini vengono indistintamente impiegati per fare riferimento alla totalità degli esemplari appartenenti alla categoria di oggetti studiati.

In geografia, i processi di definizione dei campioni sono più articolati. La scelta del campione infatti, riguarda non solo la popolazione da sottoporre ad indagini, ma anche la porzione di spazio geografico da analizzare. Essendo quello dello stanziamento umano un fenomeno che vede l'interazione della specie umana con variabili e fattori spaziali, la campionatura deve prevedere sia gli insediamenti in questione, sia il territorio, oltre che le loro caratteristiche distintive. Ad esempio, immaginando lo studio della distribuzione di strutture agrarie, si dovrà tenere conto delle caratteristiche dei suoli oltre che della morfologia. Così, l'assenza di fattorie all'interno di un territorio di montagna sarà meno significativa dell'assenza delle stesse in un territorio pianeggiante con ottimi suoli per la coltivazione.

Nella ricerca geografica, lo studio di fenomeni sociali all'interno di un'area specifica – qui intesa come campione spaziale – deve tenere conto non solo del campione stesso, ma anche delle variazioni delle caratteristiche fisiche e culturali del territorio in questione.

4.4.1 Campionamento della popolazione

La domanda più ricorrente fra gli studenti che si avvicinano per la prima volta ai metodi di quantificazione spaziale è "qual è il numero minimo di osservazioni per considerare un campione sufficientemente rappresentativo?" Ovvero, quale sia il numero minimo di misurazioni da raccogliere per avere la certezza che la loro analisi potrà produrre informazioni veritiere o che rispecchino in qualche misura le caratteristiche complessive della popolazione.

Nel campo della geografia umana, il campione solitamente si riferisce a forme di stanziamento umano. L'uso dei suoli fa riferimento prevalentemente al tema della campionatura del territorio. Cercando di rispondere alla domanda

di partenza, potremmo affermare che non esiste una soglia minima. In realtà la dimensione del "campione ideale" è legata all'estensione e all'intensità del fenomeno. Se l'oggetto di studio è la città etrusca, allora il campione sarà per forza di cose molto ridotto. Si tratta di un campione chiuso, nel senso che il totale delle città corrisponde a 10 unità dato che questo valore rappresenta la totalità. Se invece il campione sarà la fattoria del XIX secolo, allora è chiaro che il campione dovrà essere molto più ampio; inoltre esso sarà parziale, nel senso che per definizione non corrisponderà alla totalità delle fattorie ma solo ad una parte di esse.

Come regola generale si può intuire dunque che, tenendo conto dell'intensità del fenomeno, maggiore sarà il campione, maggiore sarà il suo livello di significatività. Perciò, nel caso delle scienze umane, ed in particolare delle discipline storiche, per le quali si possiede, nella maggior parte dei contesti di ricerca, di una serie parziale dei dati possibili, il problema non è tanto misurare a priori la quota di rilevamenti da raggiungere, quanto, una volta finita la raccolta dei dati, riuscire a valutarli e a cogliere il loro livello di significatività.

Sotto questo punto di vista, appare chiaro che il problema centrale sia, nella prima parte della pubblicazione dei risultati, la descrizione delle procedure adottate e del metodo con il quale il campione dei dati studiati è stato definito.

Un'ultima considerazione relativa alla consistenza dei campioni. Nel campo dell'analisi delle caratteristiche distributive dei fenomeni spaziali occorre ricordare come spesso lo stesso campione possa essere per certi tipi d'analisi sufficiente, mentre per altri no. Come è già stato suggerito sopra, bisogna sempre avere a mente il fatto che le misurazioni si riferiscono agli attributi degli oggetti e non agli oggetti stessi. Così, è l'attributo a definire quale sia il migliore metodo per misurare la natura stessa del campione minimo indispensabile per raggiungere risultati significativi.

4.4.2 Campionamento del territorio

Nel processo di raccolta dei dati e delle informazioni a livello spaziale, è fondamentale comprendere la distinzione tra la raccolta casuale dei dati e la campionatura ordinata.

La raccolta casuale si basa sul principio secondo il quale il miglior modo per definire il campione significativo in un deteminato territorio sia quello di raccogliere i dati liberamente senza uno schema predefinito o strategia di ricognizione. In sostanza, si raccolgono i dati dove capita. Questo principio prevede a sua volta, in termini geografici, che vi sia un elevato grado di "visibilità" degli oggetti studiati; fattore che comporta una certa abbondanza di

informazioni disponibili. La raccolta casuale si basa anche sul fatto che il ricercatore tratta l'intera popolazione come un campione unico nel quale non si desidera realizzare nessuna forma di discriminazione sulla base di attributi; in qualche misura mira all'identificazione di elementi distintivi generali senza nessuno schema locazionale predefinito[23].

Va considerato però che l'utilizzo di questa strategia presuppone la raccolta delle informazioni in modo omogeneo lungo l'intera superficie sotto esame. In altre parole, la libertà implicita nel processo di raccolta casuale non deve assegnare maggiore peso al rilievo in una determinata area piuttosto che in un'altra. Proprio per questo motivo, tale metodo ha una forte valenza nel processo di identificazione dell'andamento generale in termini spaziali degli attributi analizzati sull'intero territorio.

Contrariamente alla raccolta casuale, vi è il cosiddetto campionamento ad aree; chiamato anche campione arbitrario, nel senso che il ricercatore definisce arbitrariamente il perimetro dove andare a raccogliere le informazioni. Questo tipo d'approccio risponde alla necessità di indagare il territorio con maggiore dettaglio, spesso proprio per un problema legato alla visibilità o scarsità dei dati disponibili. Fattore che costringe a ricognire un intero territorio con grande intensità.

Oltre a questo, il campionamento ad aree conferisce maggiore importanza o peso al ruolo delle localizzazioni (intese come posizioni) rispetto a quello casuale. In altre parole, questo approccio può essere utilizzato per indagini che mirano non tanto all'identificazione dei valori generali in un determinato territorio, quanto alla differenziazione specifica delle singole aree sotto esame. Le aree campione divengono zone caratteristiche per le quali si stabiliscono così dei valori di correlazione spaziale con gli attributi osservati. In parole semplici, il campionamento ad aree tende, per sua natura, a identificare le differenze e non tanto gli aspetti generali.

Il principale limite di questo tipo di approccio è che, nel processo di scelta delle aree campione, il ricercatore potrebbe selezionare aree caratterizzate o da concentrazioni limite o da vuoti in termini di frequenza degli oggetti studiati. Questo comporterebbe una visione falsata o distorta dell'area di studio. Nello stesso modo, processo di aggregazioni e vuoto potrebbero non essere identificati quando si procede alla selezione delle aree campione.

Il campionamento ad aree, in certe circostanze, può naturalmente essere integrato dall'utilizzo dalla raccolta casuale di informazioni.

[23] Holmes 1967.

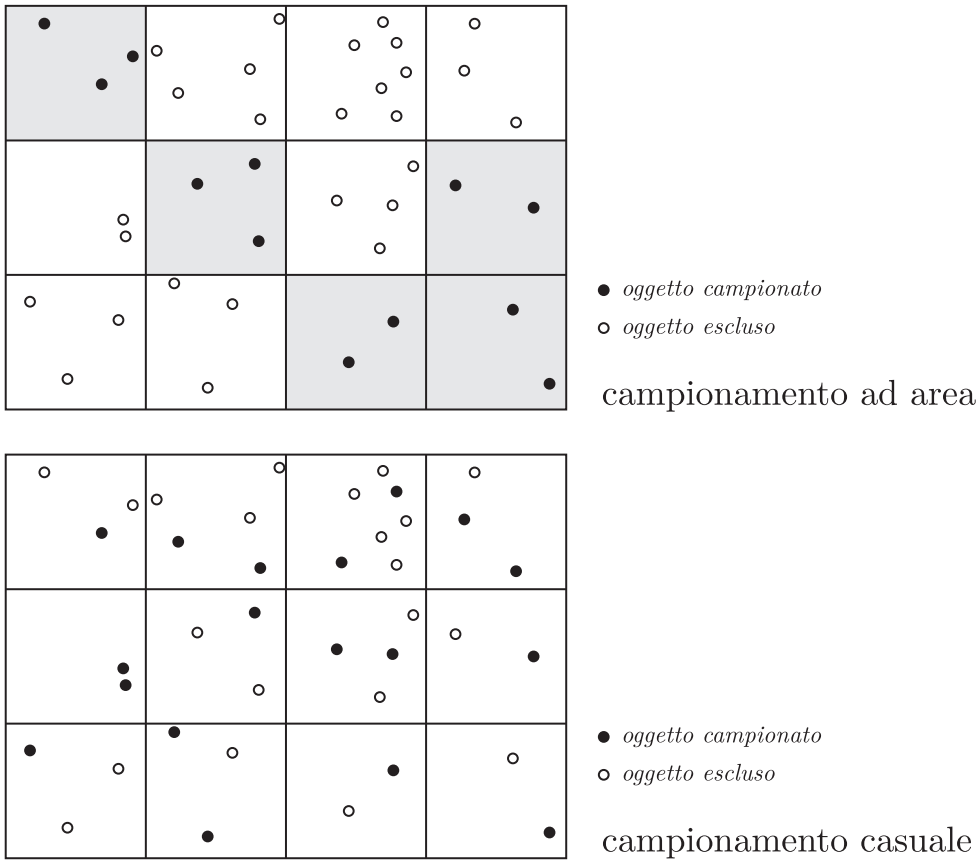


Fig. 4.3. Distinzione tra un campionamento territoriale e un processo di acquisizione casuale delle informazioni.

4.5 Statistica descrittiva

Come è già stato spiegato, l'analisi spaziale può essere intesa, nel suo complesso, «come un caso particolare dell'analisi dei dati»[24]. Gran parte dei metodi in questa disciplina sono basati sull'applicazione dei metodi statistici ed in particolare di quelli della cosiddetta "statistica descrittiva". Quest'ultima, come d'altronde suggerisce il suo nome, ha come scopo quello di riepilogare le caratteristiche generali, in termini di somiglianza (misure di tendenza centrale) o differenziazione (misure di dispersione) di una variabile.

[24] Landini, Massimi 1986, p. 13.

Come verrà illustrato nei seguenti paragrafi, proprio per la sua funzione di sintesi generale, la statistica descrittiva è fondamentale per la produzione cartografica in ambito scientifico. In altre parole, la conoscenza – anche elementare – dei principi della statistica descrittiva è essenziale per la produzione di mappe relative ai fenomeni analizzati.

Ogni processo di quantificazione formale ha inizio con la raccolta di un gruppo di misurazioni sotto forma di variabile. Si prenda in considerazione, ad esempio, la tabella 4.1. Essa riporta diverse informazioni relative ad alcuni comuni della provincia di Siena. Ogni colonna rappresenta una variabile.

Un primo aspetto che si può osservare nella tabella 4.1 è l'abbondanza di informazioni. Essa appare densa di numeri che spesso si è soliti associare con una ricchezza di nozioni. Fin qui niente di particolare, dato che la rappresentazione dei dati sotto forma di tabelle complesse caratterizza la maggior parte della produzione scientifica di alcuni contesti disciplinari delle scienze umane.

Pur nell'abbondanza di dati a disposizione, un altro aspetto che salta subito all'occhio del lettore, è la difficoltà di cogliere la sintesi delle informazioni illustrate[25]. Scopo della statistica descrittiva è proprio quello di cogliere tale essenza. In altri termini, la statistica descrittiva può essere considerata come l'approccio metodologico con il quale è possibile giungere alla sintesi o semplificazione di grandi insiemi o gruppi di dati.

La sintesi dei dati non va intesa come un semplice o banale processo di riduzione del volume complessivo delle informazioni. In realtà, diversi ambiti della ricerca fanno uso della statistica descrittiva per il semplice motivo che è proprio nella sintesi e classificazione dei dati che si può cogliere a pieno il potenziale d'informazione latente nei dati registrati.

4.6 La classificazione dei dati e le distribuzioni di frequenze

Se è vero che gran parte delle osservazioni geografiche può essere codificata sotto forma di variabili all'interno di una cornice formale, è altrettanto vero che tali misurazioni devono essere per prima cosa ordinate in modo da potere essere effettivamente analizzate. Il primo e forse più importante processo di organizzazione di una variabile è la "classificazione" delle distribuzioni di frequenza. Gran parte del processo di quantificazione dipende da questa fase, che può essere dunque considerata come una prima ricognizione delle informazioni. Se il processo di classificazione verrà portato a termine nel modo corretto, ci sono molte possibilità che l'analisi o individuazione dei *trend* possano essere identificati con successo. Viceversa, senza una corretta classificazione dei dati, probabilmente si andrà a concludere in modo sterile il processo d'analisi.

[25] Silk 1979, p. 8.

				Popolazione		
n	comune	area km^2	perimetro	1951	1961	1971
1	Abbadia San Salvatore	59,1	40,9	7053	8553	8519
2	Asciano	215,8	95,7	9734	8070	5867
3	Buonconvento	64,5	55,5	4444	3775	3020
4	Casole d'Elsa	148,4	82,7	5263	4168	3023
5	Castellina in Chianti	99,8	59,3	4886	3647	2917
6	Castelnuovo Berardenga	177,0	89,5	9937	7835	5110
7	Castiglione d'Orcia	141,5	65,7	5244	4650	3573
8	Cetona	53,1	39,8	4866	4290	3397
9	Chianciano Terme	36,7	32,8	4548	5489	6788
10	Chiusdino	141,9	74,0	4824	4061	2862
11	Chiusi	58,1	40,2	8674	8848	8771
12	Colle di Val d'Elsa	92,3	56,4	12063	12884	14812
13	Gaiole in Chianti	129,1	65,9	5437	3978	2894
14	Montalcino	243,7	91,0	10203	8825	6297
15	Montepulciano	165,8	67,8	17365	15820	14356
16	Monteriggioni	99,3	59,0	5248	4386	5512
17	Monteroni d'Arbia	105,8	60,6	5593	4931	4756
18	Monticiano	109,3	63,6	2972	2343	1875
19	Murlo	114,6	56,8	3469	2523	1890
20	Piancastagnaio	69,6	40,6	5324	5583	4697
21	Pienza	122,6	80,6	4770	3960	2987
22	Poggibonsi	70,6	59,9	14387	18634	25386
23	Radda in Chianti	80,2	53,1	2932	1946	1588
24	Radicofani	118,4	63,1	2748	2288	1605
25	Radicondoli	133,0	69,6	3227	2210	1320
26	Rapolano Terme	82,9	47,0	5541	5783	5188
27	San Casciano dei Bagni	92,1	62,4	3819	3231	2461
28	San Gimignano	138,3	53,0	11297	10039	7673
29	San Giovanni d'Asso	66,4	53,3	3263	2352	1421
30	San Quirico d'Orcia	42,2	36,9	2332	2297	2261
31	Sarteano	85,1	50,5	4702	4253	3759
32	Siena	118,5	67,2	52566	61453	65634
33	Sinalunga	78,4	50,8	10838	11345	11274
34	Sovicille	143,6	68,7	7671	6498	5364
35	Torrita di Siena	58,1	41,3	6678	6469	6753
36	Trequanda	63,9	53,7	3519	2645	1611
	Totale	3819,8		277437	270062	257221

Tabella 4.1. Dati riassuntivi relativi ai comuni della provincia di Siena.

La classificazione può essere intesa come il processo di organizzazione dei dati con lo scopo di identificare le frequenze dei valori all'interno di una variabile[26]. Si tratta di un processo che ha come scopo principale quello di organizzare i dati in base a "classi" o "categorie" di appartenenza, sulla base delle informazioni relative ad ogni singola misurazione. Vi è necessità di organizzare e dare senso a una variabile è per facilitare la lettura, ma soprattutto diper assimilare il significato stesso di questi dati.

L'aspetto più importante dell'identificazione delle distribuzioni di frequenza è che questa elaborazione consente di ottenere una visione di sintesi dei dati registrati. Come verrà illustrato di seguito, classificare una distribuzione comporta la perdita dei dati specifici ma rende possibile una visione d'insieme altrimenti non percepibile dall'intero corpo dei dati. Ad esempio, benché sia precisa e tutto sommato chiara, la tabella 4.1 difficilmente si presta ad una illustrazione ed una interpretazione semplificata dei dati in essa raccolti. Essa è in grado di registrare il dato specifico, ma non di rilevare le tendenze generali all'intero del corpo delle informazioni.

Data una certa variabile, ad esempio l'area dei comuni della provincia di Siena, è lecito interrogarsi su come questi valori siano distribuiti all'interno di un intervallo definito da un limite inferiore e da uno superiore. Per distribuzione si intende proprio la frequenza in cui i valori saranno registrati all'interno di ogni classe. Per classe s'intende la definizione formale di uno di questi gruppi di misurazione. Così, ad esempio, nel caso dell'estensione dei comuni le classi possono essere: i comuni con un'area tra 30 e 51 km^2; i comuni con un'area tra 52 e 73 km^2 e così via di seguito. Il risultato finale della classificazione sarà dunque il conteggio di quanti valori cadono – o si registrano – all'interno di ogni singola classe[27]. In altri termini, il calcolo della frequenza per ognuna di esse[28].

L'importanza della classificazione statistica in geografia va ben al di là dello stesso processo di quantificazione. Di fatto, e come è stato appena suggerito, la classificazione comporta l'introduzione del concetto di classe che costituisce uno degli strumenti più importanti a livello cartografico sia in fase di gestione che di analisi dei dati. Non solo: il processo di analisi di una distribuzione di frequenze attraverso la classificazione di una determinata variabile costituisce un passaggio essenziale per fare emergere il reale potenziale dell'informazione

[26] Bohrnstedt, Knoke 1998, pp. 39-40.

[27] Per una trattazione dettagliata del problema della definizione degli intervalli di classe in campo geografico si rimanda a Evans 1977, pp. 98 e ss.

[28] Il processo di classificazione di una variabile deve rispettare il principio di disgiunzione; ovvero un determinato valore o misurazione può appartenere a una sola classe. In altri termini, le classi, non avendo intervalli comuni, sono disgiunte.

Fig. 4.4. Carta dei comuni della provincia di Siena.

osservazioni	$5 \times (\log_{10}(n))$
10	5
30	7.3
50	8.4
100	10
250	11.9
500	13.4
1000	15
2000	16.5
3000	17.3

Tabella 4.2. Anche se non può essere intesa come una regola fissa, spesso si fa riferimento alla funzione $5 \times (\log_{10}(n))$ per stabilire il numero massimo di classi sulla base delle n osservazioni.

al suo interno. Come al solito, il linguaggio comune non rende giustizia alla semplicità e al valore dei concetti. Perciò procediamo con una dimostrazione.

Immaginiamo dunque di dovere rappresentare la distribuzione di frequenze relative all'estensione dei diversi territori comunali della provincia di Siena. Il primo passo nel processo di classificazione è quello di definire le classi. Di fatto, senza una previa fase per una loro definizione, non vi può essere processo di classificazione.

Prendendo in considerazione i dati osservati per l'estensione dei comuni, si può osservare che i dati si collocano all'interno di un intervallo generale che va da 36 a 243 km^2. Ciò significa, più semplicemente, che non ci sono comuni con un'estensione inferiore a 36 km^2 né superiore a 243 km^2. Dunque l'intervallo 36-243 rappresenta lo spettro che verrà suddiviso grazie alla definizione delle classi. Questo intervallo è comunemente definito "campo di variazione" o "intervallo di variazione". Nel caso dell'area dei comuni, il campo di variazione equivale a:

$$243 - 36 = 207 km^2$$

4.6.1 La costruzione delle classi

Il primo passo per compiere la classificazione dei dati è quello di suddividere il campo di variazione in un numero di classi che si adatti al tipo di dati raccolti. In effetti non esiste una procedura unica per la definizione delle

classi[29]. In generale, i due fattori che determinano la scelta della loro quantità sono il numero di misurazioni e l'ampiezza del campo di variazioni. Pur non essendo una regola fissa, un intervallo di variazione può in prima istanza essere scomposto in un numero che si colloca normalmente tra 7 o 12 classi. Non è in ogni caso molto comune osservare distribuzioni di frequenze con meno di 5 e più di 16 classi. Questo perché nello studio di una distribuzione non è molto utile sintetizzare oltre misura con poche classi, nè perdere di vista le tendenze generali dando troppo spazio agli aspetti particolari con la definizione di molte classi. Un altro principio utilizzato frequentemente è quello di non fare più classi di $5 \times (\log_{10}(n))$, dove n rappresenta il numero di osservazioni[30].

Solitamente, quando si fa un'analisi esplorativa dei dati, è richiesta la costruzione di classi ad intervalli regolari. Questo significa che l'estensione dell'intervallo di ogni classe dovrà essere equivalente.

$$\text{intervallo delle classi} = \frac{\text{campo di variazione}}{\text{numero delle classi}}$$

In questo esempio verranno definite 7 classi; per comodità, nel processo di suddivisione, l'intervallo di variazione verrà ampliato leggermente tra 35 e 245 km^2; dunque un campo di variazione di 210. Ogni classe avrà in questo modo un intervallo di 30 unità. La prima sarà definita dall'intervallo 35-64, la seconda dall'intervallo 65-94 e così di seguito. Nel caso della prima classe 35 e 64 vengono a costituire i suoi "limiti". Suddividendo l'intervallo di variazione in 7 parti di uguale ampiezza, si ottengono le seguenti classi:

classe	km^2	numero di comuni
1	35-64	8
2	65-94	9
3	95-124	8
4	125-154	7
5	155-184	2
6	185-214	0
7	215-244	2
	totale	36

Tabella 4.3. Distribuzione delle superfici dei comuni della provincia di Siena.

[29] Evans 1977, p. 100: «Deciding on the number of classes is a subjective stage... »
[30] Silk 1979, p. 12.

Una volta definite le classi si procede semplicemente alla conta. Alla fine del processo, ogni classe descriverà la frequenza di comuni al suo interno. In altre parole, 8 comuni hanno un'estensione compresa tra i 35 e i 64 km^2, 9 comuni nell'intervallo tra 65 e i 94 km^2 e così via (vedi tabella 4.3, terza colonna).

Va tenuto conto del fatto che gli istogrammi possono essere costruiti anche con classi a intervalli di diversa lunghezza. Ovvero, non è necessario che gli intervalli siano tutti di uguale lunghezza. In molti casi questa soluzione si presta per raggruppare i dati nella coda (ultime classi dell'istogramma).

4.6.2 Confini e valore centrale delle classi

Come illustrato nella tabella 4.3, il nome di una classe, ad esempio 35-64, viene utilizzato per identificare l'intervallo specifico di quella stessa. In linea teorica, questa classe dovrebbe contenere tutti i numeri compresi tra 35 e 64; questa affermazione non è però molto precisa. Dato che le classi sono separate da limiti superiori e inferiori – rappresentate per altro da numeri interi consecutivi – è necessario chiarire esattamente in quale punto viene a collocarsi invece il confine fra le due classi[31].

La classe 65-94 ha come limite inferiore 64.500 km^2 e come limite superiore 94.500 km^2. Questo significa che se un comune presenta un estensione di 64.526 km^2, esso dovrebbe essere registrato nella seconda classe della tabella 4.3; se invece la sua superficie fosse di 64.449 km^2 resterebbe nella prima classe.

Un modo concreto per identificare il confine di una classe è quello di sommare semplicemente il limite superiore ed inferiore di due classi consecutive e poi dividere per due:

$$\text{confine tra I e II classe} = \frac{64 + 65}{2} = 64.500$$

Naturalmente il problema dell'identificazione dei confini passa in secondo luogo nel caso dell'individuazione di distribuzioni di frequenze di variabili discrete. Al di là di tutte queste considerazioni, si tenga presente anche che di solito non ci si deve preoccupare della precisione dei limiti delle classi dato che generalmente sono gli stessi *software* ad occuparsi di assegnare in modo automatico i valori sulla base dei confini estremi delle classi definite.

Oltre ai confini, ogni classe possiede un suo valore centrale. Questo è calcolato facendo la somma dei limiti estremi divisa per 2. Nel caso della prima classe della tabella 4.3:

[31] Confini e limiti di una classe non vanno confusi.

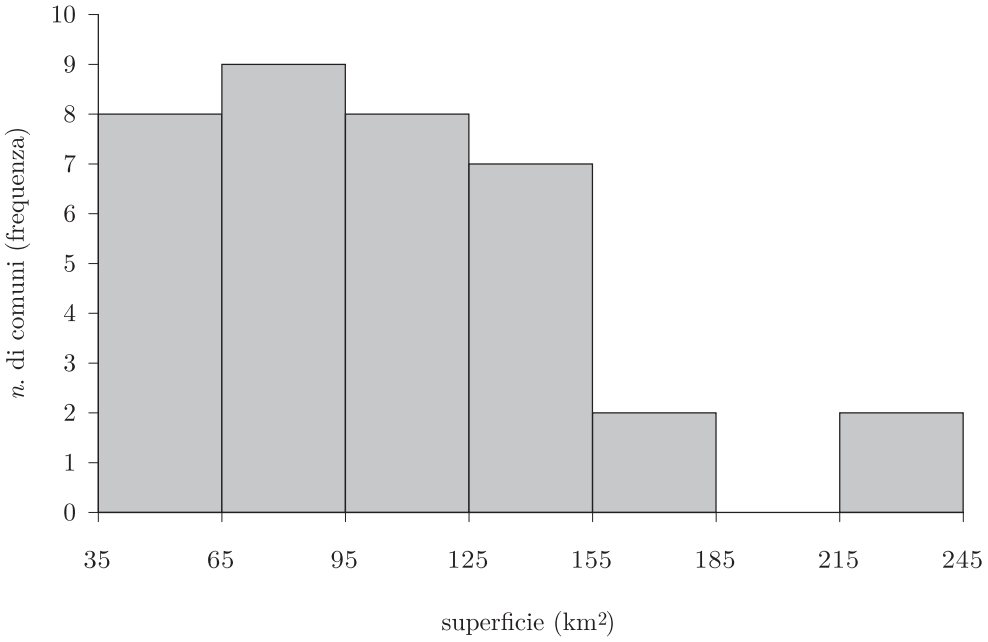


Fig. 4.5. Istogramma dei valori relativi alla superficie dei comuni della provincia di Siena.

$$\text{valore centrale I classe} = \frac{35 + 64}{2} = 49.5$$

Così la prima classe 35-64 ha come valore centrale 49.5. In base a quanto affermato prima, l'intervallo della classe è 30. Ciò significa che 15 unità si trovano sopra e 15 sotto a tale valore. Questi intervalli (superiore ed inferiore) dovrebbero corrispondere poi ai confini appena segnalati. Così, nel caso della prima classe:

$$\text{confine superiore I classe} = 49.5 + 15 = 64.5$$

4.6.3 Istogrammi di frequenza

Come spiegato sopra, l'operazione di identificazione della frequenza di distribuzione di una variabile sotto forma di tabella (vedi tabella 4.3) permette

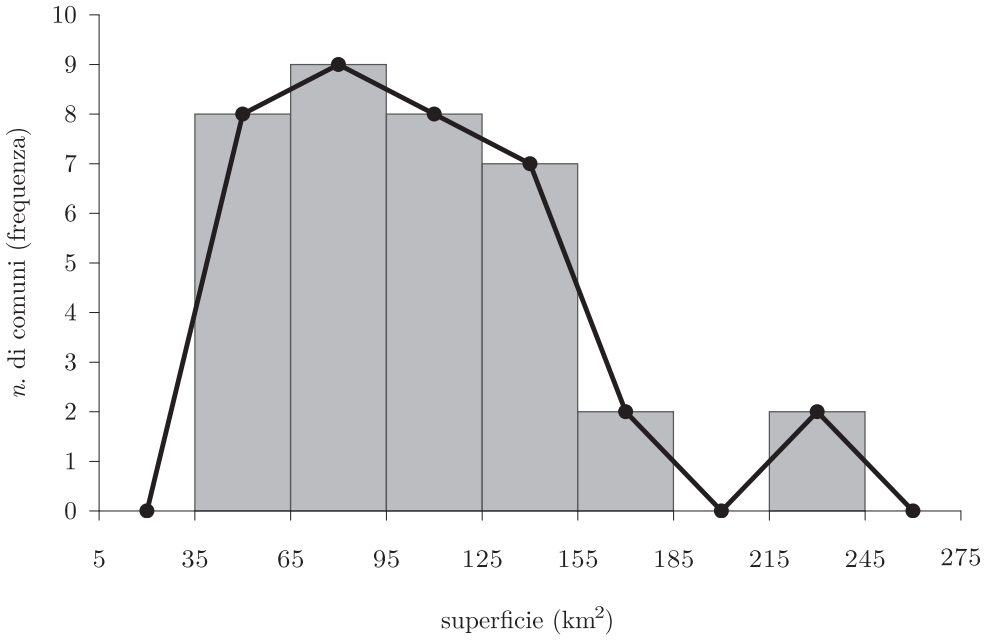


Fig. 4.6. Poligono della distribuzione di frequenza delle superfici dei comuni della provincia di Siena.

di cogliere con maggiore chiarezza che nel caso dei dati grezzi (tabella 4.1) la distribuzione delle aree dei comuni lungo tutto il campo di variazione. In modo da facilitare ulteriormente la lettura di questa sintesi si ricorre spesso alla realizzazione di un grafico speciale conosciuto come "istogramma".

L'istogramma è un grafico a barre utilizzato per illustrare i risultati di un processo di classificazione. La costruzione di un istogramma è un processo sostanzialmente semplice, ma che deve rispettare regole specifiche. Un aspetto fondamentale della statistica descrittiva è che la rappresentazione dei dati sotto forma di tabelle o grafici non è un aspetto discrezionali. Esistono delle regole e convenzioni ben specifiche che fanno sì che certe scelte, benché gradevoli a livello grafico, siano del tutto sbagliate a livello formale. In questo i *software* commerciali più diffusi non sono di grande aiuto.

L'istogramma deve rappresentare sulle ordinate il valore di frequenza (numero di casi) mentre sulle ascisse le classi. Così, ogni singola barra è la rappresentazione grafica della frequenza di osservazioni che ricadono nell'intervallo di ogni singola classe.

Le ordinate rappresenteranno sostanzialmente una scala su cui registrare le frequenze. L'ascissa invece sarà suddivisa in tante parti quante sono le classi di

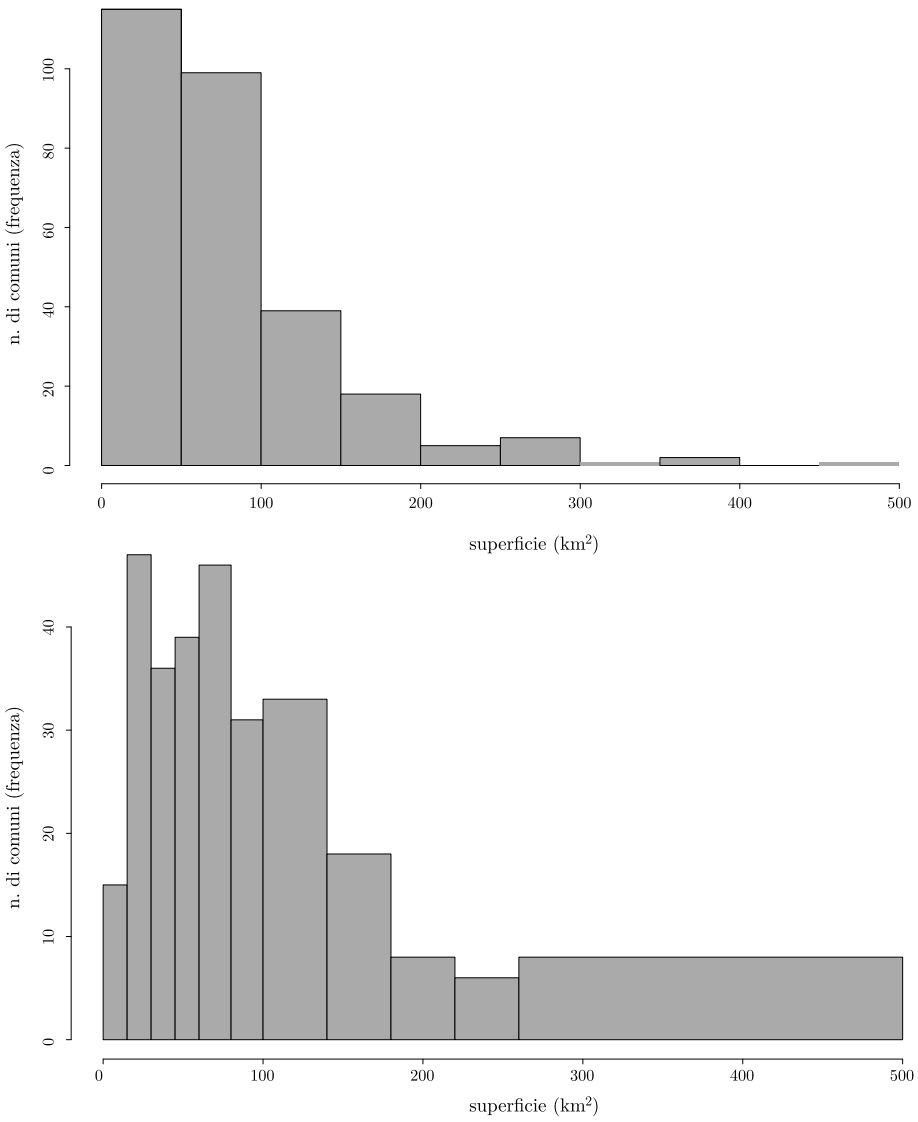


Fig. 4.7. La costruzione delle classi non deve necessariamente seguire una suddivisione ad intervalli regolari. In molti casi i grafici possono illustrare meglio la distribuzione dei dati con classi non omogenee. Nell'immagine sono illustrati gli istogrammi delle superfici dei comuni della Toscana. Si tratta di un campione molto ampio caratterizzato da una grande variabilità. Come si può osservare, l'istogramma irregolare (basso) illustra meglio la distribuzione dei valori.

frequenza. La scala numerica dovrebbe essere indicata con i valori centrali di ogni singola classe. Questa soluzione consente di evitare uno spreco di spazio. Nel caso dell'esempio trattato, la prima classe sarebbe dunque indicata col valore 49.5. In alternativa, per ogni barra dovrebbero essere segnalati entrambi i limiti, appunto, come nella figura, sempre per la prima classe, i valori 35-64.

Oltre all'istogramma tradizionale, è possibile rappresentare la frequenza di un determinato fenomeno con i cosiddetti grafici a poligono. Questi non differiscono nella sostanza dagli istogrammi come quelli della figura 4.5: si basano su grafici a linee che congiungono (sulle ascisse) il valore centrale di ogni singola classe all'altezza della frequenza relativa. Il risultato è un grafico a linee – o spezzata poligonale – che disegna il profilo superiore di un poligono la cui area corrisponde alla somma di quelle delle barre di un grafico a istogrammi. La figura 4.9 illustra chiaramente il concetto grazie alla sovrapposizione dell'istogramma (nello sfondo) con il grafico a poligono. Si può osservare bene come i vertici del poligono – secondo i principi di costruzione di questa tipologia di grafici – corrispondano proprio al valore centrale[32].

4.6.4 Classificazione cartografica

Il processo di classificazione ha delle applicazioni particolari in campo cartografico. Con le dovute cautele si può affermare che, se lo scopo della statistica descrittiva è quello di comprendere ed interpretare corpi di misurazioni formali, quello della geografia quantitativa è quello di estendere tale principio a livello spaziale.

Di fatto, la stessa logica utilizzata per la realizzazione di una classificazione può essere estesa alla comprensione, applicazione e costruzione di una carta geografica. In sostanza, invece di contare o raggruppare le informazioni per classi di appartenenza, la classificazione cartografica ha come scopo quello di formulare una convenzione basata su classi definite da una simbologia cartografica. Un esempio di questo concetto è la carta 4.8, su cui sono state create 6 classi di riferimento. Ognuna di queste rappresenta un intervallo relativo alle estensioni già illustrate dalla tabella 4.1.

La figura 4.8 fornisce dunque un'altra possibilità di lettura delle informazioni raccolte nella tabella 4.1, così come dell'istogramma della figura 4.5. Lo stesso tipo d'informazione acquista nuove forme di rappresentazione così come nuovi canali di assimilazione. Perciò la mappa 4.8 e l'istogramma 4.5 rappresentano in qualche misura modelli dei dati raccolti nella tabella 4.1.

Il vantaggio fondamentale della classificazione sotto forma di rappresentazione cartografica è che, oltre a costituire un sommario riassuntivo di in-

[32] Spiegel 1994, pp. 38-39; Silk 1979, pp. 12-13.

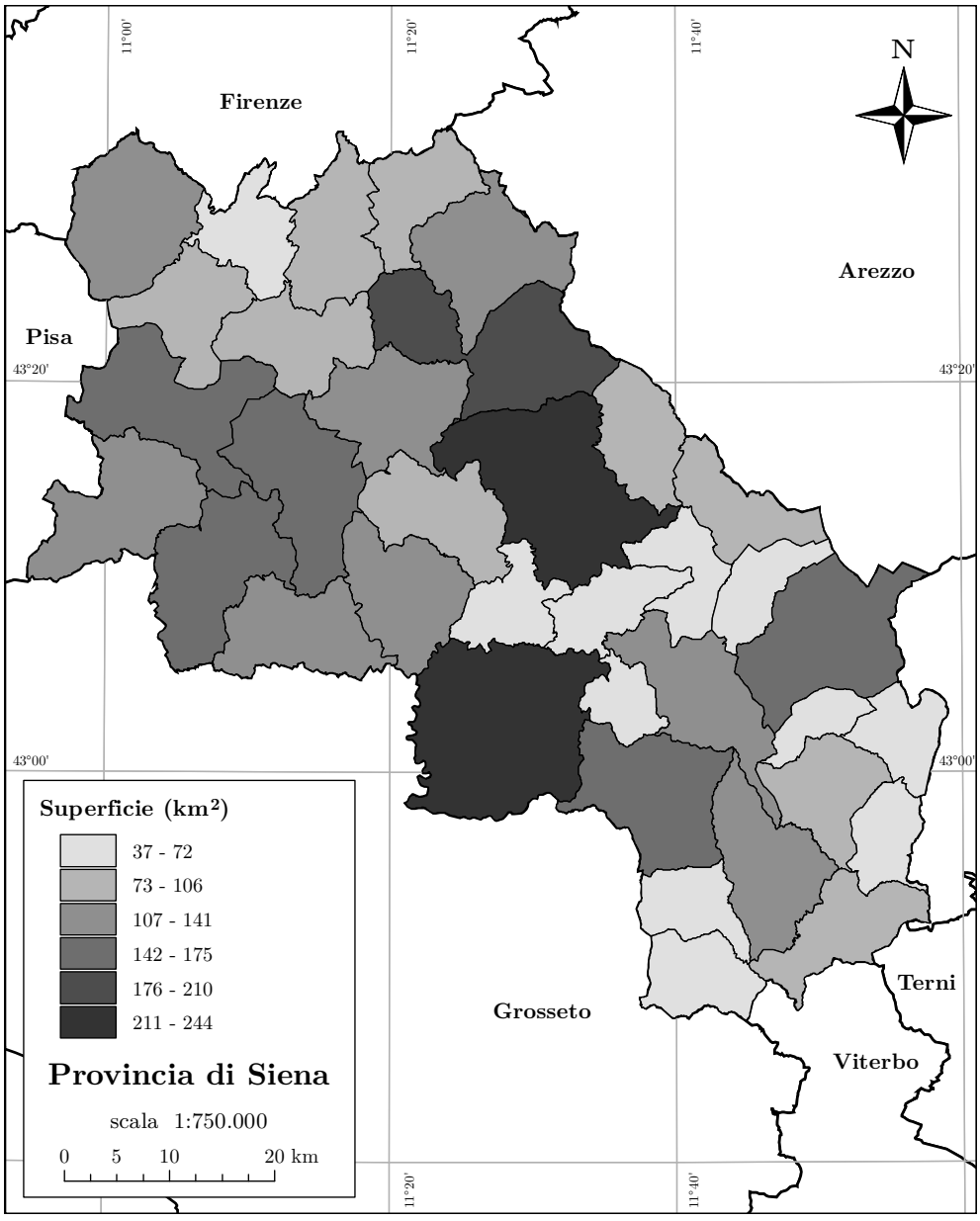


Fig. 4.8. Carta dei comuni della provincia di Siena con la rappresentazione simbolica della densità organizzata in sei classi.

formazioni grezze, essa permette di visualizzare la distribuzione delle classi in termini spaziali. L'istogramma costituisce una rappresentazione sotto un'unica dimensione, peraltro astratta. Il documento cartografico permette di osservare e cogliere sia aspetti legati alla distribuzione delle classi, sia, come in questo caso, l'estensione o la contiguità spaziale delle medesime.

4.6.5 Curve cumulative

Un altro modo di illustrare le distribuzioni di frequenza sono le curve cumulative o "ogive". Queste possono essere descritte come un grafico a poligono in cui vengono rappresentati in termini di percentuale (sul totale della popolazione) i valori cumulativi di frequenza delle classi in un ordine progressivo. In altri termini, l'ogiva rappresenta, per ogni classe, la somma delle percentuali relative ad essa e alle precedenti. Così, la prima classe corrisponde alla percentuale della sola prima classe, la seconda alla somma della percentuale della prima più quella della seconda classe, la terza alla percentuale della prima più quelle della seconda e della terza, e così via fino all'ultima, che rappresenta la somma totale delle percentuali, ovvero il 100%.

Il calcolo delle percentuali è molto elementare e si basa su una semplice proporzione. Ad esempio, per il calcolo della percentuale sul totale della prima classe, assumiamo per vero il principio secondo il quale:

$$\begin{aligned} 8 &\longrightarrow x \\ 36 &\longrightarrow 100 \end{aligned}$$

Leggeremo *8 sta ad x come 36 sta a 100*. Dunque sappiamo che 36 è il totale (100%) dei comuni. A che percentuale corrispondono allora 8 comuni? Applicando la logica delle proporzioni possiamo ricavare che:

$$\frac{36}{8} = \frac{100}{x}$$

ovvero

$$x = \frac{8 \times 100}{36} = 22.2\%$$

Una volta calcolate le percentuali per ogni singola classe, si procede a sommarle progressivamente in modo da identificare il valore cumulativo registrato nella quinta colonna della tabella 4.4.

Lo scopo della frequenza cumulativa è l'analisi, oltre che la rappresentazione, della percentuale dei casi all'interno delle classi sotto una certa soglia.

classe	km^2	numero di comuni	percentuale	percentuale cumulativa
1	35-64	8	22.2%	22.2%
2	65-94	9	25%	47.2%
3	95-124	8	22.2%	69.4%
4	125-154	7	19.4%	88.8%
5	155-184	2	5.5%	94.4%
6	185-214	0	0%	94.4%
7	215-244	2	5.5%	100%
	Totale	36	100.0%	

Tabella 4.4. Distribuzione delle aree dei comuni della provincia di Siena.

Così, ad esempio, nella tabella 4.4 si può osservare come il 69.4% dei comuni abbia una superficie inferiore a 125 km^2.

Il grafico 4.4 è molto utile, dato che permettere di rilevare in modo chiaro, lungo la sequenza delle classi, l'incremento progressivo dei valori percentuali accumulati fino a quel punto. Così, ad esempio, osservando il grafico, è possibile rilevare con estrema facilità come l'88.8% dei comuni abbia una superficie inferiore a 155 km^2.

4.7 Misure di "tendenza centrale"

Come già detto sopra, scopo della statistica descrittiva è quello di sintetizzare le caratteristiche generali di una data distribuzione. Ogni distribuzione di frequenza è caratterizzata da due categorie di indici o indicatori che ne descrivono la struttura: misure di tendenza centrale e misure di dispersione.

Il primo tipo di quantificazione di cui si occupa la statistica descrittiva sono proprio le misure di "tendenza centrale". Quando si dispone di una serie di osservazioni o misurazioni correttamente formalizzate in una variabile è possibile compiere delle operazioni di quantificazione in grado di identificare misure che riassumano e descrivano in modo sintetico le caratteristiche generali dei dati studiati. Esse vengono definite appunto misure di tendenza centrale o "indici di posizione". Le principali misure di tendenza centrale sono: la media, la moda, la mediana e il *midrange*. Questo tipo di misurazioni si basa sul principio che esiste un punto verso il quale una determinata variabile converge; tale punto costituisce un indicatore intrinseco del valore della variabile. Si potrebbe dire che questo tipo di quantificazione, in qualche misura riassuma l'essenza della variabile attraverso l'identificazione del punto intorno al quale

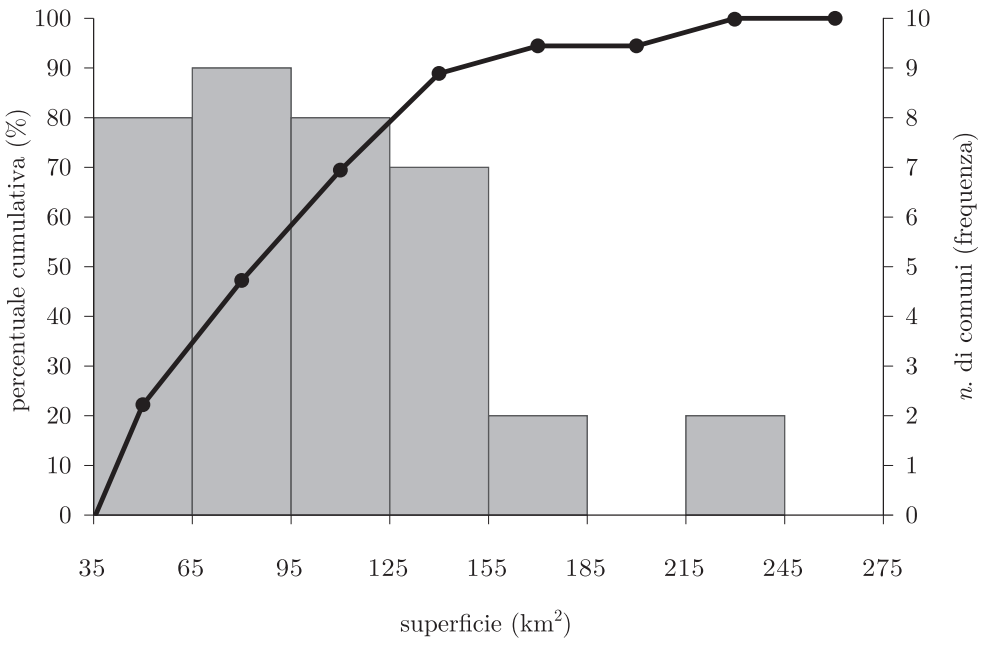


Fig. 4.9. Curva cumulativa della distribuzione di frequenza della superficie dei comuni della provincia di Siena.

i valori della medesima tendono a convergere. Ad esempio, l'estensione ideale verso la quale il territorio dei comuni di una certa provincia tende a convergere, oppure il numero medio di abitanti verso il quale tendono gli insediamenti appartenenti ad un certo rango in un territorio.

4.7.1 La media

La "media aritmetica" è la misura di tendenza centrale più comune. Essa non è altro che la somma (o sommatoria) di un insieme di valori, diviso per il numero (n) degli stessi. La media di una variabile qualsiasi, ad esempio X, viene rappresentata con il simbolo, $\bar{X}$, denominato solitamente X "soprassegnato". Così ad esempio, la media dell'area dei comuni della Provincia di Siena viene calcolata attraverso la seguente procedura:

$$\bar{X} = \frac{131 + 385 + 103.9 + \cdots + 85.9}{36}$$

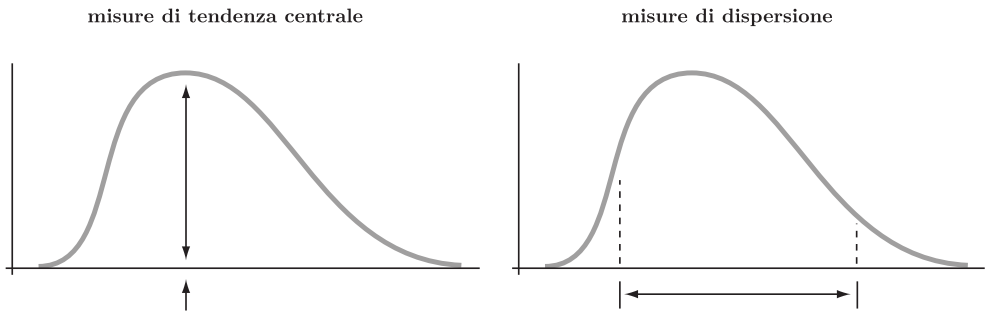


Fig. 4.10. La statistica descrittiva si occupa dell'identificazione e dell'analisi delle caratteristiche fondamentali delle distribuzioni di frequenza. Queste misure possono essere distinte in due categorie fondamentali: misure di tendenza centrale e misure di dispersione.

La media diviene in questo modo un valore di sintesi che descrive la tendenza generale, nel caso riportato qui come esempio, quello delle dimensioni dei comuni in unità d'area.

Generalmente la media viene rappresentata con l'utilizzo di una notazione particolare. Per descriverla a livello formale si ricorre all'espressione

$$\bar{X} = \frac{\sum_{j=1}^{n} X_j}{n} \tag{4.1}$$

Nell'espressione compare il simbolo Σ che viene impiegato per rappresentare la "sommatoria". Questa viene corredata da due ulteriori indicazioni: in primo luogo in basso $j = 1$ che sta ad indicare che questa sommatoria riguarda la somma di tutti i valori a partire dal primo; il secondo elemento è il simbolo n in alto che sta ad indicare che la sommatoria riguarda tutti gli n valori della variabile X_j. La cosa che occorre ricordare è che la notazione classica non è altro che un modo "grazioso" di scrivere quanto detto prima.

$$\text{sommatoria dei valori della variabile } X_j = \sum_{j=1}^{n} X_j$$

che significa

$$\sum_{j=1}^{n} X_j = X_1 + X_2 + X_3 + \cdots + X_n$$

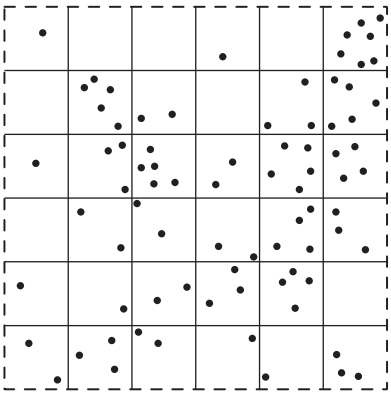


Fig. 4.11. Distribuzione per punti all'interno di una griglia. Una volta classificata, la frequenza media per cella può essere calcolata con l'aiuto della media ponderata.

Dunque, la media della variabile X si calcola come segue:

$$\bar{X} = \frac{\sum_{j=1}^{n} X_j}{n}$$

Da un punto di vista puramente geografico si può immaginare che le dimensioni dei comuni non siano un attributo arbitrario, ma seguano un *trend* definito dai rapporti spaziali definiti dall'interazione storica dell'uomo con lo spazio e dalla progressiva conquista da parte dei centri maggiori dei propri territori di pertinenza. In altri termini, le loro aree restano tendenzialmente – sia nella pratica che nella teoria – entro limiti concreti. Si può immaginare, ad esempio, che un comune di 2 km^2 sia un'evenienza molto rara. Lo stesso varrebbe per un comune con dimensioni di 5000 km^2. Si può anche immaginare che esista, almeno per quest'area geografica, una dimensione ideale. In poche parole, esiste una tendenza generale che fa sì che le aree dei comuni di questa regione convergano attorno ad un certo valore centrale.

Una volta calcolata la media di A otteniamo dunque come risultato per $\bar{A} = 82,5\ km^2$. Ciò significa che l'insieme di valori registrati per questa variabile tende a convergere attorno a tale valore. Rappresenta una sintesi generale dell'estensione dei comuni della provincia; ovvero, un valore astratto e rappresentativo verso il quale tutte le misurazioni registrate in una determinata variabile tendono a convergere.

Oltre a quella aritmetica, esiste anche la "media ponderata". Questo tipo di misurazione viene impiegato per il calcolo della media di valori previamente

frequenza punti	*numero di celle*
0	7
1	7
2	8
3	6
4	3
5	4
6	0
7	1

Tabella 4.5. Classificazione della frequenza dei punti all'interno delle singole celle della figura 4.11.

raggruppati (come ad esempio in qualche processo di classificazione) e possiede un peso per ogni gruppo o classe. Va calcolata con la seguente formula:

$$\bar{M} = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i} \tag{4.2}$$

dove x_i rappresenta la sequenza delle classi, e w_i le rispettive frequenze. La formula può essere semplicemente intesa come:

$$\text{media ponderata} = \frac{\text{somma di tutte la classi moltiplicate per il rispettivo peso}}{\text{somma di tutti i valori di frequenza}}$$

Si prenda in considerazione la figura 4.11. La frequenza dei punti all'interno delle celle è stata sintetizzata nella tabella 4.5. È stato rilevato che 7 celle hanno 0 punti al loro interno; in 7 celle si registra 1 punto e così di seguito. Qualora si volesse calcolare la frequenza media dei punti all'interno di ogni singola cella a partire dai dati della tabella 4.5 sarebbe sufficiente applicare la formula 4.2 nel seguente modo:

$$\bar{M} = \frac{(0 \times 7) + (1 \times 7) + (2 \times 8) + \cdots + (7 \times 1)}{7 + 7 + 8 + 6 + 3 + 4 + 0 + 1} = 2.22$$

4.7.2 La mediana

Anche la "mediana" rappresenta un indice di posizione o di tendenza centrale. Questa misura indica, per un dato campione di misurazioni, il valore

relativo alla posizione centrale. Solitamente la mediana viene indicata con il simbolo M_e. Un altro modo per concepire la mediana è quello di immaginarla come la posizione o il punto che divide l'intera popolazione in due parti uguali. Così metà della popolazione o delle misurazioni resterà sotto tale valore, mentre l'altra metà sopra.

Per calcolare la mediana, per prima cosa si deve procedere a mettere in ordine crescente i valori del campione.

Così, ad esempio, nel caso la variabile F, sono stati registrati i valori

$$41, 15, 67, 16, 23, 11, 40$$

e si procede ad ordinarli in modo ascendente:

$$11, 15, 16, 23, 40, 41, 67$$

In questo caso la mediana corrisponde alla quarta posizione perché è quella che divide esattamente in due parti (di tre posizioni ciascuna) la distribuzione. Nel caso la distribuzione sia composta da un numero dispari di valori, si può utilizzare la seguente formula per calcolare la posizione della mediana:

$$M_e = x_{\frac{n+1}{2}} \tag{4.3}$$

Così, nell'esempio precedente

$$M_e = x_{\frac{7+1}{2}} = x_4 = 23$$

Perciò, l'identificazione della posizione della mediana all'interno di una serie di dati dispari, non presenta particolari problemi. Nel caso di una serie pari, invece, dove non esiste un unico valore centrale, è necessario calcolare la media aritmetica dei due valori di mezzo. Ad esempio, la mediana dei numeri

$$11, 15, 19, 25, 26, 32$$

può essere calcolata con la seguente formula:

$$M_e = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} \tag{4.4}$$

La media dei due valori centrali risulterà dunque 22, vale a dire: (19+25)/2. Ad esempio, per ottenere la mediana dei dati relativi all'area dei comuni della provincia di Siena, è necessario innanzitutto procedere a disporre la serie delle misurazioni in ordine ascendente. In questo caso, partendo dall'estensione del comune di Chianciano Terme (36.7 km^2), San Quirico d'Orcia (42.2 km^2), ecc. fino ad arrivare a Montalcino (243.7 km^2). La sequenza ordinata dei comuni appare dunque la seguente:

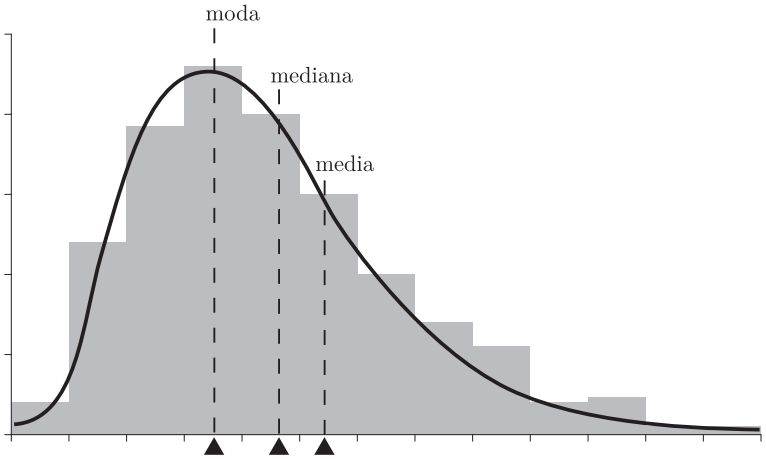


Fig. 4.12. Rappresentazione grafica delle tre principali misure di tendenza centrale.

36,7 42,2 53,1 58,1 58,1 59,1 63,9 64,5 66,4 69,6 70,6 78,4 80,2 82,9 85,1
92,1 92,3 99,3 99,8 105,8 109,3 114,6 118,4 118,5 122,6 129,1 133,0 138,3
141,5 141,9 143,6 148,4 165,8 177,0 215,8 243,7

Successivamente si contano le posizioni, fino a trovare quella rispetto a cui, la metà delle misurazioni risulta a sinistra e l'altra metà a destra. Nell'esempio si dispone di un totale di 36 misurazioni. Di conseguenza non esiste un vero valore centrale e perciò si rende necessario l'applicazione di 4.4.

$$M_e = \frac{x_{\frac{36}{2}} + x_{\frac{36}{2}+1}}{2} = \frac{x_{18} + x_{19}}{2} = \frac{99.3 + 99.8}{2} = 99.55 km^2$$

Nel caso dei comuni della tabella 4.1 la mediana è dunque 99.55 km^2.

4.7.3 La moda

Nello studio delle misure di tendenza centrale la "moda" di una distribuzione è intesa come il valore di maggiore frequenza.

Ad esempio, prendendo in considerazione la seguente distribuzione

$$11, 11, 15, 19, 19, 19, 19, 25, 26, 26, 32$$

la moda risulta essere 19, dato che, semplicemente, si tratta del valore registrato più volte.

In questo caso si dice che la distribuzione è "unimodale", vale a dire che possiede una sola moda. D'altra parte è possibile anche registrare una distribuzione con più di una moda, come nel seguente caso:

$$11, 15, 19, 19, 19, 25, 26, 26, 26, 32$$

dove sia 19 che 26 devono essere considerate come mode della distribuzione. In questo caso si dice che la distribuzione è di tipo "bimodale", ovvero possiede due mode.

La moda è un tipo di indice immediatamente assimilabile a variabili di tipo discreto. Mentre invece per calcolare la moda di una variabile continua è necessario ricorrere ad un istogramma in modo da identificare la "classe modale" ovvero quella che semplicemente registra la maggiore frequenza.

4.7.4 Il *midrange*

Il "*midrange*" è anch'esso una misura di posizione. Questo indice di tendenza centrale rappresenta la media tra il valore più basso e quello più alto di una serie. Così, il calcolo del *midrange* della variabile x può essere eseguito con la seguente formula:

$$\frac{x_{min} + x_{max}}{2} \tag{4.5}$$

Il *midrange* rappresenta il punto centrale del campo di variazione. È fortemente condizionato dai valori estremi e per questo motivo raramente viene utilizzato per indicare o misurare le tendenze centrali di una distribuzione.

4.8 Misure di dispersione

Abbiamo già visto come le misure di posizione forniscano un'indicazione delle tendenze centrali di una variabile o di un gruppo di misurazioni. Benché le misure di posizione possano essere indicative delle qualità complessive di una variabile, è anche vero che queste da sole possono offrire una descrizione solo parziale delle caratteristiche di una popolazione[33]. Questo perché, anche se i valori tendono a raggrupparsi intorno ad un punto (da ciò il nome di misura di posizione) è anche vero che i valori tendono a disporsi in modo più o meno disperso, e in modo più o meno intenso all'interno del campo di variazione della distribuzione. Da qui la necessità di rilevare, oltre alle misure di tendenza centrale, le cosiddette misure di variabilità o dispersione che, insieme alle prime, costituiscono le statistiche basiche di una variabile.

[33] Matthews 1985, p. 42 e seg.

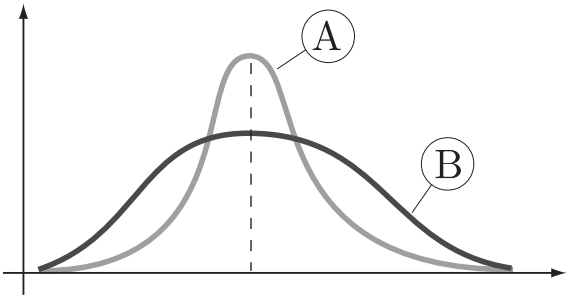


Fig. 4.13. Nel grafico sono rappresentate due distribuzioni di frequenza. Le curve possiedono una media ed una mediana (linea tratteggiata) identiche. Pur condividendo la stessa tendenza centrale le due curve sono diverse. Tale discrepanza è determinata dalla differenza nella dispersione delle frequenze. Nella distribuzione A i valori sono raggruppati maggiormente al centro. Nel caso della distribuzione B, invece, i dati sembrano disporsi in modo più omogeneo lungo l'intero campo di variazione. Dunque l'indice di "dispersione" della distribuzione B è superiore a quello di A.

La ragione per le quale le misure di posizione non sono sufficienti per capire le caratteristiche di una certa popolazione, può essere compresa con il seguente esempio. Si immagini una popolazione di 100 individui. Per capire meglio le caratteristiche del gruppo è stata calcolata la media dell'età in 25 anni. Può questo valore descrive accuratamente le caratteristiche della popolazione indagata? Successivamente si procede ad analizzare questa popolazione più nel dettaglio e si scopre che il gruppo di 100 individui è composto da 50 individui di 10 anni e 50 di 40 anni. L'età media di questo gruppo, 25 anni, non rappresenta dunque un indice caratteristico di nessuno dei membri che lo compongono. In questo caso la media viene a costituirsi perciò come un dato sviante sulla popolazione studiata. Se l'esempio appena proposto può apparire come un caso estremo, è anche vero che, in nessun caso, la media o le altre misure di tendenza centrale possono fornire dati esaurienti e definitivi per descrivere una certa distribuzione.

Se, oltre alla misura di posizione si forniscono altre indicazioni relative alla dispersone dei valori, ci si potrebbe fare un'idea molto più chiara del tipo di popolazione analizzata. Le misure di dispersione di una variabile sono il campo di variazione, i quartili, la varianza e lo scarto quadratico medio.

4.8.1 Campo di variazione

Il "campo di variazione" o *range* è la più semplice delle misure di dispersione di una variabile. Questo valore si ottiene per una variabile calcolando la differenza tra il valore più alto e quello più basso. Può essere definito come l'intervallo più piccolo che contiene tutte le osservazioni registrate. Anche se tale misura può dare delle indicazioni generali a proposito del campo entro il quale i dati sono dispersi, essa non indica in nessun modo come i valori siano dispersi al suo interno. Il campo di variazione di una variabile x va calcolato dunque con:

$$\text{campo di variazione} = x_{max} - x_{min} \tag{4.6}$$

L'unità del campo di variazione è la stessa delle osservazioni registrate nella variabile analizzata. Come già spiegato in precedenza, la posizione centrale del campo di variazione corrisponde al *midrange*.

Nel caso della superficie dei comuni della provincia di Siena, il campo di variazione equivale a

$$\text{campo di variazione} = 243.7 - 36.7 = 207 km^2$$

Il che significa che la differenza tra la superficie maggiore registrata per i comuni e quella minore è pari a 207 km^2.

4.8.2 I quartili

Un altro modo intuitivo per individuare il grado di dispersione di una distribuzione di frequenza è quella di definire i "quartili". Questi sono i tre valori (Q_1, Q_2, Q_3) che dividono una distribuzione ordinata esattamente in quattro parti.

Il primo quartile (Q_1), conosciuto anche come quartile inferiore, corrisponde al valore posto nel punto che divide il primo 25% delle misurazioni dalla parte restante.

Il secondo quartile (Q_2) coincide con la mediana ($Q_2 = M_e$). In pratica il secondo quartile divide la distribuzione in due insiemi di dati, ognuno dei quali con il 50% delle misurazioni (vedi 4.7.2).

Il terzo quartile (Q_3), infine, equivale al valore che si colloca nel punto che divide la distribuzione tra un primo 75% dei dati e il restante 25% delle misurazioni.

Benché molto elementare come principio, i quartili, insieme al campo di variazione, forniscono uno schema empirico per determinare le caratteristiche di una distribuzione.

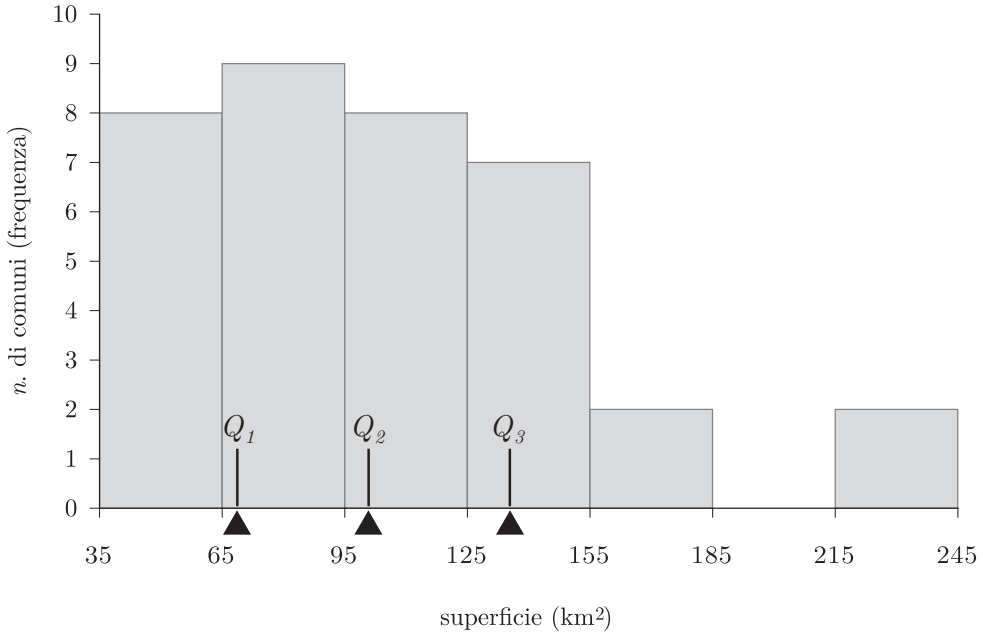


Fig. 4.14. Rappresentazione grafica dei quartili in relazione all'istogramma dei valori relativi alla superficie dei comuni della provincia di Siena: $Q_1 = 68.7$, $Q_2 = 99.5$, e $Q_3 = 134.1$.

Nel caso la posizione dei quartili non corrisponda ad un valore specifico, ma bensì ad un punto a metà tra due numeri, si procede con il computo della loro media. Ad esempio, nel caso di una variabile con i valori 5, 9, 11, 21, 23, 29, 35, 36, 39, 40, 45, i rispettivi quartili sono:

$$\begin{aligned} Q_1 &= 16 \\ Q_2 &= 29 \\ Q_3 &= 37.5 \end{aligned}$$

Questo perché

$$\begin{aligned} Q_1 &= (11 + 21)/2 \\ Q_2 &= M_e \\ Q_3 &= (36 + 39)/2 \end{aligned}$$

4.8.3 La media interquartile

A partire dai quartili è possibile calcolare altre due misure importanti che sono la "media interquartile" e il "*range* interquartile".

Le media interquartile è calcolata semplicemente come la media tra primo e terzo quartile. Questa naturalmente costituisce una misura di tendenza centrale. La sua caratteristica principale è che non viene influenzata dai valori estremi o *outliners* e per questo motivo è preferita in molti casi al *midrange*. Essa viene calcolata semplicemente come:

$$\text{media interquartile} = \frac{Q_1 + Q_3}{2} \tag{4.7}$$

Il "*range* interquartile" o IQR invece presenta l'intervallo relativo al campo di variazione coperto tra Q_1 e Q_3. Perciò viene calcolato con:

$$\text{IQR} = Q_3 - Q_1 \tag{4.8}$$

4.8.4 La varianza

Esiste invece tutta una serie di indici di dispersione che ha per scopo quello di rilevare come i valori di una distribuzione di frequenza siano sparsi all'interno del campo di variazione. Gli indici maggiormente usati nell'ambito della statistica descrittiva sono la "varianza" e lo "scarto quadratico medio" o "deviazione standard". Come verrà illustrato di seguito, entrambe queste misure sono collegate.

Il calcolo della dispersione fa solitamente riferimento ad un punto, o soglia, attorno al quale i valori tendono a disperdersi. Naturalmente, all'interno del campo di variazione, il luogo naturale all'interno del quale i valori tendono a disporsi è la media. Per questo motivo, si può affermare che le misure di dispersione indicano sostanzialmente il grado di dispersione di una certa variabile dalla propria media.

Ad esempio, se si prendano in considerazione le due seguenti distribuzioni:

$$125, 128, 133, 133, 134, 139, 140, 145$$

$$23, 46, 77, 133, 151, 171, 195, 215$$

si può osservare a occhio nudo come queste differiscano notevolmente in termini di dispersione attorno alla propria media. Le misure che verranno trattate di seguito hanno come scopo quello di rilevare in modo formale tale dispersione.

La prima misura è la varianza, solitamente rappresentata con il simbolo σ^2 per indicare quella di un'intera popolazione ("varianza della popolazione") oppure con il simbolo s^2 nel caso si tratti della varianza di un campione estratto da una popolazione ("varianza campionaria").

La varianza rappresenta in sostanza "la media dei quadrati degli scarti dei valori di una data variabile x dalla media". In realtà il concetto è molto più semplice di quanto le parole lascino intuire. La formula della varianza è la seguente:

$$\sigma^2 = \frac{\sum_{j=1}^{n}(x_j - \bar{x})^2}{n} \tag{4.9}$$

Come si può osservare, la varianza corrisponde ad una media delle differenze al quadrato dalla media ($\bar{x}$). In sostanza si sottrae ogni singolo valore dalla media e lo si eleva poi al quadrato: $(x_j - \bar{x})^2$. Così la varianza indica la dispersione media attorno alla media. Il risultato del calcolo della varianza non è espresso nell'unità di misura delle misurazioni originali della variabile, bensì nel loro quadrato. Tuttavia, dato che i valori sono elevati al quadrato, la varianza rappresenta le differenze complessive (positive e negative) sulla media.

Molto simile alla varianza è la varianza campionaria che date le proprietà specifiche viene usata più di frequente ed in particolare nel caso di distribuzioni non molto grandi:

$$s^2 = \frac{\sum_{j=1}^{n}(x_j - \bar{x})^2}{n-1} \tag{4.10}$$

Per capire meglio questo concetto si prendano in considerazione i valori:

$$123, 57, 97, 141, 104, 27$$

Cercando di calcolare la varianza campionaria s^2 a partire da 4.10 si ottiene dunque:

$$\begin{aligned} s^2 &= \frac{(123 - 91.5)^2 + (57 - 91.5)^2 + \cdots + (27 - 91.5)^2}{6 - 1} \\ s^2 &= 1795.9 \end{aligned}$$

Il che sta ad indicare un elevato livello di dispersione. Ma dato che tale misura è espressa approssimativamente come il quadrato della media delle

differenze dalla media della distribuzione il suo valore non è immediatamente intuibile. Per renderlo più chiaro e in molti casi utile al fine della descrizione di una distribuzione si fa ricorso allo scarto quadratico medio.

4.8.5 Lo scarto quadratico medio

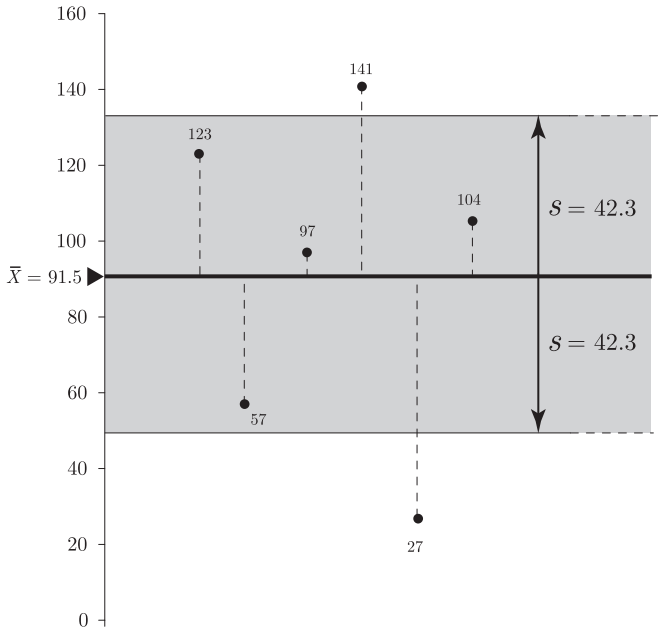


Fig. 4.15. Rappresentazione grafica dello scarto quadratico medio (S). Lo scarto quadratico medio può essere inteso come la media delle differenze di una variabile dalla propria media. In questo caso, 42.3 rappresenta la media degli scarti dalla media: (91.5) linea nera orizzontale al centro del grafico.

Lo scarto quadratico medio (s) rappresenta la media dello scarto di una determinata variabile dalla sua media. Questo valore si ottiene semplicemente calcolando per una data variabile la radice quadrata aritmetica della varianza campionaria. Infatti l'equazione dello scarto quadratico medio è:

$$s = \sqrt{\frac{\sum_{j=1}^{n}(x_j - \bar{x})^2}{n-1}} \tag{4.11}$$

Il principale vantaggio del calcolo dello scarto quadratico medio è che i valori sono riferibili alle unità di misura della variabile. Dunque, questi sono immediatamente assimilabili e forniscono un'indicatore molto utile relativo alle caratteristiche della distribuzione.

Prendendo i valori 123, 57, 97, 141, 104, 27 si può calcolare lo scarto quadratico medio in

$$s = \sqrt{\frac{(123 - 91.5)^2 + (57 - 91.5)^2 + \cdots + (27 - 91.5)^2}{6 - 1}} = 42.3$$

Si prenda in considerazione la superficie in km^2 di tutti i comuni della regione Toscana. Partendo da una base GIS ed estrapolando i dati dalla tabella del relativo piano cartografico, si ottengono i valori della superficie dei comuni:

182,5 48,8 73,7 73,9 55,7 53,9 29,3 62,4 180,6 62,2 50,4 272,0 19,4 42,1
60,5 34,8 48,5 26,2 57,3 26,7 11,4 109,1 28,7 27,1 19,9 31,0 44,0 66,5 41,3
154,1 77,2 24,1 85,5 33,1 70,6 63,4 53,5 94,2 133,3 236,3 95,6 31,3 79,2
49,4 146,3 80,1 26,4 47,8 28,0 138,8 16,6 14,5 79,2 61,7 99,2 36,6 32,1
43,0 34,2 30,8 33,4 8,9 85,3 77,0 97,3 57,0 16,0 16,0 42,1 17,7 11,7 68,2
7,8 59,2 12,7 114,4 45,9 32,6 31,4 48,8 42,2 45,9 15,7 25,0 28,8 7,2 29,6
22,3 28,6 38,4 102,3 6,0 86,5 54,6 54,1 18,8 65,2 83,3 25,9 53,5 49,3 74,1
80,0 59,7 103,9 156,1 24,9 42,9 54,2 121,7 24,6 38,0 11,6 62,2 48,7 102,5
26,9 36,3 78,6 107,8 66,5 125,0 26,4 14,9 59,5 169,4 14,9 37,6 30,2 18,4
86,5 90,9 46,2 66,5 77,7 71,5 59,9 73,7 19,7 131,0 47,5 44,8 104,7 27,1 85,9
42,4 22,7 57,0 21,4 75,3 60,9 25,5 83,6 66,0 92,5 385,0 43,3 56,7 19,3 80,5
36,6 99,8 138,2 24,0 29,4 70,6 46,5 128,8 131,5 100,1 72,6 62,0 252,7 92,3
177,0 58,8 99,7 154,9 111,4 89,9 343,0 148,6 118,5 23,8 83,0 253,2 20,2
44,8 215,7 143,6 23,6 40,8 14,3 78,6 65,6 132,5 105,9 98,1 142,0 64,0 63,5
58,2 64,8 165,4 108,3 66,5 26,5 102,6 122,9 243,9 42,1 58,1 36,6 141,7 84,9
118,1 53,5 67,7 49,4 59,0 92,1 93,3 63,5 69,7 64,3 174,5 81,7 48,3 164,3
72,1 70,5 36,5 185,8 156,1 62,8 75,5 18,1 82,6 113,9 92,0 67,4 29,5 23,0
25,2 186,2 108,0 101,8 66,6 11,6 120,6 45,6 42,9 141,6 114,6 109,5 283,5
92,7 33,3 192,9 284,4 83,3 19,3 129,9 163,9 56,0 161,6 88,3 162,2 473,6
209,4 19,8 124,8 16,6 37,7 273,5 6,0 45,3 13,3 39,8 45,8 251,0 371,9 101,9
227,1 187,4 60,5 21,6

Segue dunque una sintesi generale grazie alla classificazione e ad alcune misure di tendenza centrale e di dispersione:

$$
\begin{aligned}
\bar{X} &= 80.0 \\
M_e &= 62 \\
Q_1 &= 33.7 \\
Q_2 &= M_e \\
Q_3 &= 100.9 \\
\text{media interquartile} &= 67.3 \\
S^2 &= 4582.6 \\
S &= 67.7
\end{aligned}
$$

4.9 Le distribuzioni statistiche

È stato già illustrato come gran parte delle possibili osservazioni nel campo della geografia umana possano essere codificate all'interno di una cornice formale di riferimento. Misurazioni che possono essere registrate sotto forma di variabili grazie all'utilizzo di diversi tipi di scale di misurazione (vedi paragrafo 4.2). È stato anche spiegato come, organizzando e classificando questi dati, sia possibile identificarne la frequenza all'interno di un campo di variazione che permette di rilevare aspetti molto importanti di un certo fenomeno (paragrafo 4.6). Infine è stato introdotto il tema della statistica descrittiva e si è detto come questa fornisca degli strumenti essenziali, in grado di sintetizzare in un modo sufficientemente accurato gruppi di osservazioni.

Lungo i paragrafi precedenti, si è fatto ricorso al concetto di distribuzione; è stato infatti utilizzato più volte come sinonimo di frequenza e in alcuni grafici è stata usata una curva continua come rappresentazione simbolica dell'andamento degli istogrammi di frequenza. In questa sezione si cercherà di definire con maggiore precisione e dettaglio questo concetto, che costituisce, fra le altre cose, uno degli strumenti di maggiore importanza nell'ambito della statistica.

Il tema della distribuzioni statistiche appartiene alla "statistica inferenziale", che può essere intesa come il settore che, attraverso il computo delle probabilità, si occupa dello studio delle caratteristiche di un determinato fenomeno o sistema[34].

[34] La statistica va distinta in statistica descrittiva e statistica inferenziale.

4.9.1 Probabilità e insediamento

Una "distribuzione" o "distribuzione statistica" (da non confondere con le distribuzioni spaziali) costituisce un'astrazione formale delle probabilità che un certo evento arbitrario (o un certo gruppo di eventi arbitrari) si verifichi. Questo tipo di evento viene definito solitamente come "variabile casuale". Si prenda in considerazione il lancio di una moneta: in questo caso si sa che di norma (cioè senza condizionamenti e influenze esterne) le probabilità che si registri l'evento "testa" sono 50% e quello "croce" 50%. In campo statistico tali aspettative vengono solitamente rappresentate come 0.5 per testa (t) e 0.5 per croce (c). Questo punto è molto importante e occorre soffermarsi un istante per comprenderlo meglio, dato che si tratta della scala con la quale la scienza in generale misura le probabilità.

Solitamente 0 viene inteso come il non verificarsi di un evento X, 1 è inteso come il verificarsi dell'evento X. Nel caso dell'evento X, le probabilità che si verifichi vengono espresse con $P(X) = 1$. Nel caso del lancio della moneta le probabilità che si registri "testa" sono 0.5[35]:

$$P(X = t) = \frac{1}{2} = 0.5$$

dato che sono due le possibilità di uscita: o "testa" o "croce". Dunque potremmo dire che il calcolo delle probabilità può essere realizzato con[36]:

$$P(X = t) = \frac{n \text{ risultati favorevoli al verificarsi dell'evento}}{\text{numero di possibili combinazioni dei risultati}} \tag{4.12}$$

Il concetto di distribuzione è in qualche misura legato all'idea di probabilità, dato che, a partire della misurazione di qualsiasi variabile e dalla sua organizzazione all'interno di un istogramma di frequenza, si può giungere all'identificazione di una distribuzione probabilistica. Ed è qui che sorge la prima difficoltà per il lettore che si avvicina per la prima volta a queste tematiche: a cosa possono servire i calcoli di probabilità del lancio di monete, dadi o combinazioni di carte, per lo studio della realtà geografica?

Nel caso specifico del lancio della moneta, lo studio delle probabilità non permette di indovinare il risultato del lancio successivo. Al contrario, permette di pronosticare che, su 1000 lanci della moneta, si dovranno registrare tendenzialmente metà teste e metà croci. A livello pratico, può ad esempio servire

[35] Per un'introduzione generale alla teoria della probabilità si consiglia la lettura di Silk 1979, p. 35; Bohrnstedt, Knoke 1998, pp. 78-79; Matthews 1985, pp. 56-58.

[36] Si veda in particolare il concetto "approccio classico a priori" in Levine, Krehbiel, Berenson 2002, p. 133.

a capire se la moneta sia truccata. L'esempio appare molto banale e perciò inutile. Ma cosa succede se, invece del lancio delle monete, lo studio delle probabilità, inteso come studio formale delle frequenze osservate, viene esteso alla comprensione della realtà?

Benché sia chiaro che le osservazioni, e dunque le misurazioni della realtà non riguardano fenomeni prettamente arbitrari, ovvero, che una variabile geografica (o qualsiasi altra misurazione della realtà) non possa essere paragonata al concetto di variabile casuale – come ad esempio il lancio di una moneta – è vero invece che la logica delle distribuzioni statistiche, e con essa tutta la teoria delle probabilità, può essere utilizzata per studiare e comprendere in profondità le caratteristiche delle variabili reali. In altri termini, a livello campionario, le frequenze osservate per un certo tipo di variabile reale possono essere concepite come distribuzione di probabilità; qui intesa come la proporzione che caratterizza una certa classe di eventi o fenomeni all'interno di un campione classificato. Ad esempio, se per campione rappresentativo di insediamenti (80 nuclei) in un determinato territorio viene raccolta la variabile *quota sul livello del mare*, le frequenze osservate possono essere utilizzate come proporzioni da estendere al resto della popolazione (740 nuclei) e fare un pronostico di quali siano le frequenze all'interno delle singole classi per l'intera popolazione. Nel campo della geografia umana, lo studio della realtà in termini di probabilità è fondamentale soprattuto nello studio di quei settori per i quali non è possibile disporre di tutto il campione. Ad esempio, la statistica inferenziale e in particolare la teoria delle probabilità sono fondamentali in campo cartografico[37]. È il caso dell'archeologia dei paesaggi, nell'ambito della quale spesso si è costretti a lavorare con campioni molto ridotti.

Una distribuzione statistica è una funzione di tipo matematico che descrive in modo sintetico le frequenza di una determinata variabile. In altre parole, è un'equazione che riproduce una curva che in qualche misura rispecchia l'andamento di un istogramma di frequenza. La prima domanda che verrebbe spontaneo porsi è dunque quale sia il motivo per cui ci si dovrebbe preoccupare di identificare la distribuzione statistica di una data variabile. Oltre ai vantaggi che derivano nel processo di comparazione fra contesti o fenomeni, lo studio delle distribuzioni ci permette di formulare modelli più complessi e più articolati della realtà. Il processo di identificazione della funzione più conforme viene definito "modellazione"; nel senso che l'equazione viene a rappresentare un modello della frequenza osservata[38].

[37] Vedi ad esempio Wrigley 1977, pp. 138-139.

[38] Tanto importanti sono i modelli per la geografia, quanto importante è l'attività di modellazione.

È opportuno sapere che gran parte dei fenomeni reali è più regolare di quanto si possa immaginare. Esiste un ordine nascosto in ogni fenomeno o evento che ci circonda. Per ordine s'intende la regolarità che si può osservare nella forma degli istogrammi relativi alle loro misurazioni. L'identificazione della distribuzione statistica di una variabile corrisponde all'identificazione di un *pattern*. Questo sempre perché, a dispetto di quello che si potrebbe pensare, le forme che l'arbitrarietà è in grado di assumere nella realtà sono poche. Ovvero, molti fenomeni, anche molto diversi fra loro – come ad esempio il caso dell'altezza delle persone o della durata delle lampadine – possono essere modellati con distribuzioni statistiche simili.

Prima di entrare nei particolari delle distribuzioni, l'aspetto che bisogna sempre tenere presente è che la statistica ha sviluppato nel tempo le funzioni matematiche che modellano nel migliore dei modi le tendenze naturali e culturali di gran parte delle osservazioni che si possono compiere sulla realtà. Sta proprio lì l'importanza e la ragione del loro utilizzo. Non esiste modo più semplice nè strada più concreta, per studiare le frequenze che farlo con l'aiuto delle distribuzioni sviluppate dalla statistica.

A dire il vero, una distribuzione statistica è già stata presentata e discussa nei capitoli precedenti. Si tratta della legge di Zipf. Essa è stata illustrata come modello geografico, ma di fatto appartiene anche alla statistica inferenziale. Come abbiamo visto, la legge di Zipf poteva essere usata come schema per pronosticare le dimensioni degli stanziamenti. E, a partire dalle differenze dalla distribuzione attesa, per compiere dei confronti tra più campioni. In questo senso, il lettore tenga presente che le distribuzioni verranno trattate di seguito proprio nella logica dell'uso di quella di Zipf: creazione di un modello in grado di descrivere certe osservazioni, capace di fornire elementi utili alla definizione di schemi di confronto.

Le distribuzioni sono diverse e spesso raggruppate in famiglie. Le più importanti per il geografo sono:

- normale o gaussiana
- di Poisson
- binominale
- gamma

Al lettore si cercherà di fornire qualche elemento introduttivo relativamente a solo due tipi: curva normale e distribuzione di Poisson.

4.9.2 La distribuzione normale

La "distribuzione normale", conosciuta anche come "curva gaussiana", rappresenta la più nota e popolare di tutte le distribuzioni statistiche. La sua fama

dipende sostanzialmente dal fatto che questa funzione descrive con precisione molte delle variabili osservate in contesti naturali e culturali[39]. Si basa sul principio che molti fenomeni nella realtà fisica e sociale si organizzano intorno a una media ideale e tendono a convergere verso questa in modo graduale e progressivo, intensificando la frequenza di eventi man mano ci si avvicina a questo punto centrale. La forma della distribuzione è dunque quella di una campana (vedi figura 4.16, linea nera)[40].

La curva normale fu scoperta durante il primo quarto del XVIII secolo da Abraham de Moivre e pubblicata per la prima volta nel 1733. La sua fama è dovuta però alle applicazioni realizzate da Carl Friedrich Gauss (1777-1855) (uno dei più influenti matematici del nostro tempo) in campo astronomico. L'espressione della funzione viene solitamente rappresentata come:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \tag{4.13}$$

Questa funzione, anche se può apparire molto contorta, si basa sostanzialmente su due soli parametri: media (μ) e varianza (σ^2)[41]. I parametri di una certa distribuzione sono i valori che regolano la forma della curva. La μ definisce la posizione attesa del picco o punto centrale della campana sulle ascisse, mentre σ^2 determina l'apertura (o dispersione) della medesima.

Ogni tipo di distribuzione statistica ha di norma i propri parametri. Gran parte del lavoro di modellazione di un istogramma di frequenza consiste nell'identificare il giusto valore per i parametri in questione. Cambiando i parametri di una distribuzione cambia la forma della curva. Dunque bisogna considerare che una distribuzione non descrive in assoluto la frequenza di una data variabile. Piuttosto si può affermare che, identificati i giusti parametri di una distribuzione, si può utilizzare una distribuzione per modellare una variabile. Va tenuto presente che molti dei parametri di una data curva di distribuzione sono alcune delle misure di tendenza centrale o di dispersione viste nei paragrafi precedenti.

Per capire la logica della curva normale si partirà dall'esempio dell'altezza umana; un fenomeno che possiede frequenze compatibili con questa distribuzione statistica. Partendo da 60 misurazioni di un campione casuale di donne, si sono registrati i seguenti valori:

[39] Hammond, McCullagh 1975, p. 96.

[40] Per una introduzione generale alla curva normale si consiglia la lettura di Freedman, Pisani, Purves 1998, pp. 79-81.

[41] Così come σ^2 è usata per rappresentare la varianza di un'intera popolazione, μ rappresenta la media aritmetica di tutta la popolazione.

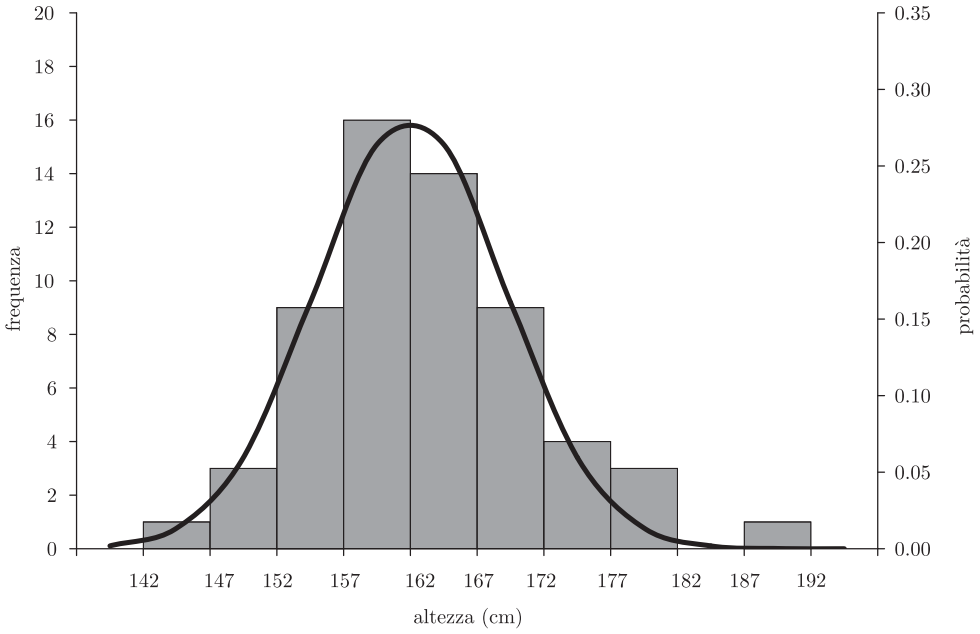


Fig. 4.16. Benché l'altezza delle persone non possa essere paragonata al lancio dei dadi, le distribuzioni statistiche possono essere di grande utilità per la descrizione di una variabile. Come si può osservare dal grafico, i valori di frequenza dell'altezza degli individui della specie umana (in questo caso femmine con $\mu = 162$ e $\sigma^2 = 50.41$) possono essere modellati sulla base della distribuzione normale: dato un campione arbitrario, non influenzato, si può fare un pronostico molto accurato dei valori che poi si andranno ad osservare.

160,1 156,7 176,9 166,1 159,7 156,5 164,7 157,7 166,8 167,6 161,0 163,1
178,7 147,7 172,6 162,8 163,5 163,8 170,5 157,0 177,8 167,4 157,4 169,5
162,4 165,7 163,7 155,4 161,3 179,5 168,8 154,5 153,5 160,6 155,6 159,7
167,7 164,2 154,2 161,7 160,2 175,6 157,0 157,3 144,6 174,2 160,5 171,5
164,3 159,2 159,0 165,8 167,2 165,1 187,1 161,8 144,9 168,5 151,5 158,9

Procedendo alla classificazione di questi valori, si sono ottenuti i dati registrati nella tabella 4.6.

Se si procede all'applicazione della funzione 4.13 per le probabilità che una data altezza si registri a partire dai parametri noti per l'altezza media delle donne ($\bar{X} = 162cm$, $\sigma^2 = 50.41$), si ottiene una curva a campana come quella illustrata nella figura 4.16. La curva è sovrapposta all'istogramma della classificazione dei dati illustrati in precedenza. Come si può osservare, la curva della

classe	frequenza
137-142	0
142-147	1
147-152	3
152-157	9
157-162	16
162-167	14
167-172	9
172-177	4
177-182	3
182-187	0
187-192	1
192-197	0

Tabella 4.6. Classificazione delle altezze di 60 donne appartenenti ad un campione casuale.

funzione con i parametri utilizzati descrive l'andamento dei valori osservati in modo molto accurato.

In geografia, la distribuzione normale è più legata a fenomeni fisici. Variabili quali dimensione degli insediamenti, superfici territoriali (quali comuni, provincie ecc.), seguono altri tipi di curve. Prendiamo ad esempio 80 misurazioni relative agli anni 1920-1999 dell'intensità di piovosità registrata da una stazione meteorologica della città di Palermo[42]:

600,1 629,4 407,5 498,3 637,3 687,4 343,1 535,7 529,9 768,7 651,8 933,6
711,2 883,4 740 520,5 510,6 549,5 480,4 821,6 677,7 407,3 550,5 521 602,1
510,3 559,4 457,5 495,9 578,5 547,5 502,7 320,6 600,3 665,9 495,1 474,6
551,4 526,6 375,6 474,8 398,5 526,4 605,9 544,5 563,8 507,1 449,6 418,3
487,1 289,8 473,6 399,9 533,6 353,3 401,5 531,5 221 535,1 531 415 409,3
544,8 506,7 388,2 466,8 420 581,5 497,5 287,7 456,5 558 442,9 483,7 461,1
588,5 857,1 592,1 547,6 521,2

[42] Micela, Granata, Iuliano 2001.

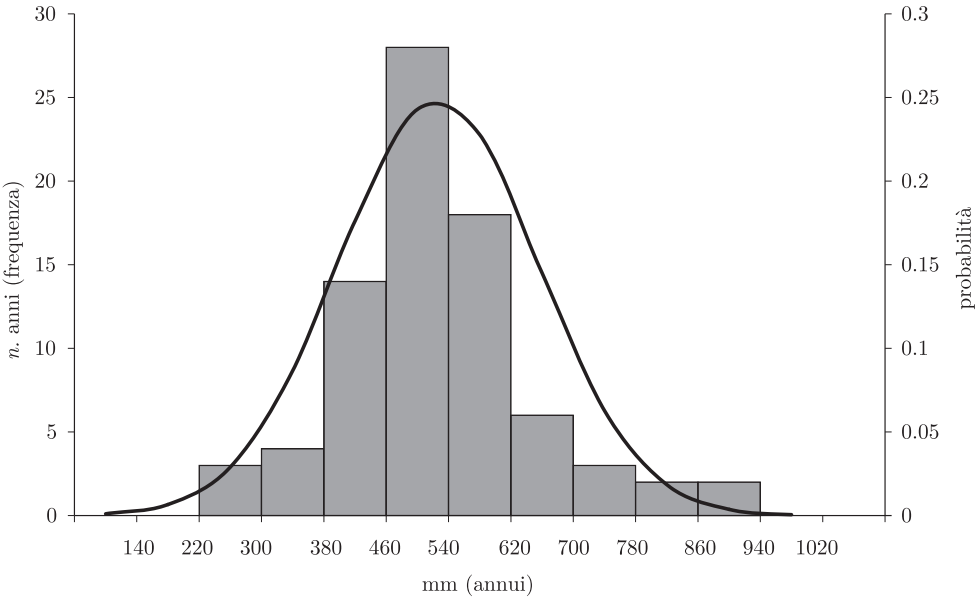


Fig. 4.17. Istogramma delle precipitazioni (mm annui) dell'osservatorio di Palermo negli anni 1920-1999, con sovrapposizione della curva normale creata con $\mu = 526.6$ e $\sigma = 16555.3$.

A partire da questi dati, si procede al calcolo della media ($\mu = 526.6$ e $\sigma = 16555.3$). Queste due misure sono utilizzate poi come parametri per la modellazione della curva[43].

Altri studi nel campo della geomorfologia hanno dimostrato la possibilità di identificare distribuzioni normali nelle pendenze registrate all'interno di territori caratterizzati da un certo livello di omogeneità[44].

4.9.3 La distribuzione di Poisson

Ai fini della comprensione delle distribuzioni geografiche e dell'analisi delle strutture spaziali, assai più utile della distribuzione normale è quella "di Poisson". Introdotta da Siméon-Denis Poisson (1781-1840) e pubblicata nel 1837,

[43] Una trattazione approfondita della conformità dei volumi delle precipitazioni con la curva gaussiana si può trovare in Wallén 1955, p. 55. Sempre in relazione all'applicazione della curva gaussiana nel campo delle precipitazioni consultare Hammond, McCullagh 1975, pp. 96-97.

[44] Strahler 1950. Altri esempi di applicazione in campo geomorfologico si trovano in Matthews 1985, pp. 31-32.

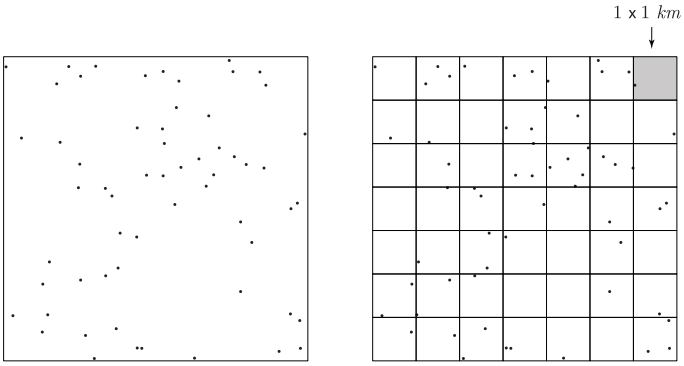


Fig. 4.18. La figura a sinistra illustra un processo casuale indipendente all'interno di un territorio di $7 \times 7\ km^2$. La figura a destra mostra la suddivisione del territorio in 49 celle di $1 \times 1\ km^2$. La distribuzione di Poisson può essere sfruttata per capire le probabilità che una cella possa contenere x punti.

questa distribuzione, che verrà trattata nel capitolo 6, rappresenta uno degli strumenti più importanti per lo studio delle carte per punti[45]. In estrema sintesi, la funzione di Poisson permette di identificare, partendo da una densità nota, la probabilità che un numero n di eventi occorra in un intervallo di tempo specifico. Tale distribuzione può essere utilizzata per capire processi arbitrari non solo su intervalli di tempo, ma anche su superfici spaziali o volumi.

Se si conosce, per un certo momento della giornata (ad esempio nel pomeriggio), la media delle persone che transitano in un determinato incrocio pedonale, la distribuzione di Poisson permette di identificare le probabilità che, in un intervallo di 1 minuto, ad esempio, passino o 0, 1, 2, o n persone. Come si può intuire subito, questo indice statistico è fortemente legato all'idea di densità e perciò fondamentale per l'analisi della carte di distribuzione.

Ma cosa ha a che vedere tutto questo con la geografia? Da tempo molti studiosi del paesaggio si sono avvalsi della distribuzione di Poisson per modellare quello che comunemente viene definito come "processo casuale indipendente". In pratica, un altro modo di chiamare una rete di distribuzioni per punti di tipo stocastico o casuale. Questo processo è caratterizzato dalle due seguenti condizioni.[46]

[45] Poisson 1837.
[46] Unwin 1986, p. 106.

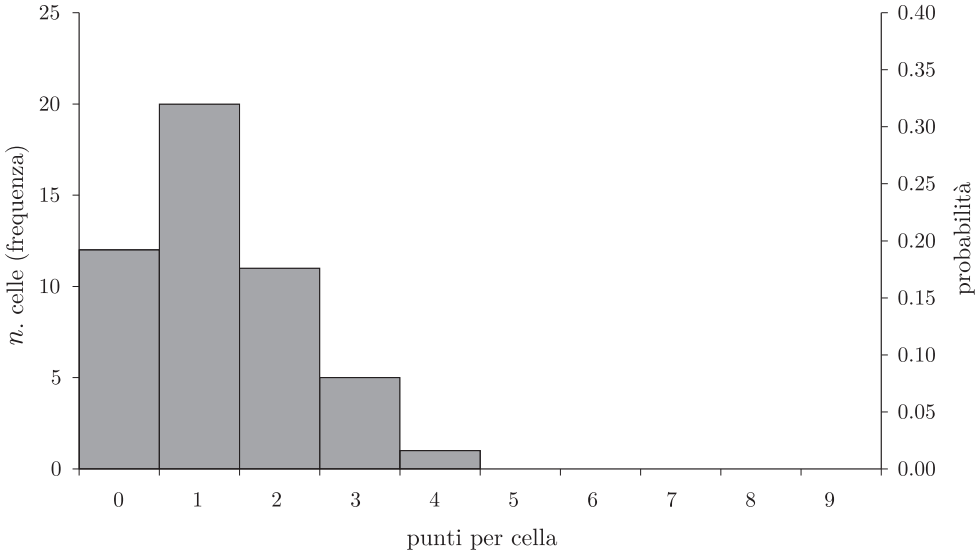


Fig. 4.19. Distribuzione di Poisson. Istogramma della frequenza dei punti per cella (da leggere con le ordinate di sinistra) e curva di probabilità calcolata con i parametri $\lambda = 1.22$ (da leggere con le ordinate di destra).

- Probabilità omogenea. In estrema sintesi, ogni area sarà caratterizzata dalle stesse probabilità di ricevere un punto. In altri termini, non ci sono zone che presentino più probabilità di altre che in esse cada un maggior numero di punti. Per molti versi la probabilità omogenea può essere paragonata alla pianura euclidea ed isotropica.
- Indipendenza. Ogni punto generato è totalmente indipendente dagli altri punti. Così gli elementi che compongono la maglia di un processo casuale indipendente non subiscono assolutamente nessun condizionamento dalla posizione di altri punti.

La distribuzione ha la seguente funzione:[47]

$$P(x) = \frac{e^{-\lambda}\lambda^x}{x!} \tag{4.14}$$

dove x rappresenta il numero di eventi (0, 1, 2 ... punti), λ la densità nota ($\lambda = n/A$) ed e la costante matematica ($e = 2.71828...$). L'espressione $x!$ indica il fattoriale del numero x. Ad esempio:

[47] Una teoria generale sull'applicazione del modello di Poisson nello studio delle carte per punti si trova in Harvey 1966, pp. 81-85.

$$\begin{aligned} 4! &= 1 \times 2 \times 3 \times 4 \\ 6! &= 1 \times 2 \times 3 \times 4 \times 5 \times 6 \end{aligned}$$

Procediamo con una dimostrazione. Si immagini un territorio di forma quadrata con lati di 7 chilometri e un'area totale di 49 km^2 (vedi figura 4.18). Si sa che in tale territorio un processo casuale indipendente ha provocato, a livello spaziale, 60 eventi registrati in una carta come punti. Suddividendo il territorio in unità o celle di 1 km^2, quali sono le probabilità che una cella registri 0 punti, 1 punto, 2 punti, ecc.?

Il primo passo è dunque calcolare λ che rappresenta la densità di punti (o valore atteso di frequenza relativo) per l'intero territorio sotto esame. Nel caso qui proposto:

$$\lambda = \frac{60}{49} = 1.22$$

Applicando la funzione 4.14 si ottengono i seguenti risultati:

x	$P(x)$	a	o
0	0.294	14.4	12
1	0.360	17.6	20
2	0.220	10.8	11
3	0.090	4.4	5
4	0.028	1.3	1
5	0.007	0.3	0
6	0.001	0.1	0
7	0.000	0.0	0
8	0.000	0.0	0
9	0.000	0.0	0
totale	1.000	49.0	49

Tabella 4.7. Nella colonna $P(x)$ sono indicate le probabilità che una cella possa contenere x punti. La colonna a riporta il numero di celle attese con x punti sulla base di $P(x)$ ($P(x) \times 49$), e infine la o mostra il numero di celle osservate con x punti. Come si può osservare dal confronto delle ultime due colonne, i valori modellati (a) e quelli effettivamente osservati(o)seguono un *trend* simile.

Risultati che confermano in larga misura quelli della prova empirica illustrata nella figura 4.18. La colonna a della tabella 4.7 riporta i valori attesi o

il prodotto di ogni singola classe x per il numero totale di celle $P(x) \times 49$, e la o i valori effettivamente rilevati dalla prova empirica. Come si può osservare dal confronto di queste due colonne, la differenza tra i valori modellati e quelli osservati è minima.

Tutto questo può apparire più o meno interessante; ed in ogni caso ci si potrebbe ancora domandare cosa abbia a che fare tutto ciò con lo studio delle caratteristiche distributive di una maglia di stanziamento. Questo esercizio serve per evidenziare come la casualità sia molto più prevedibile di quanto uno possa intuire. I processi arbitrari tendono a ripetersi in modo molto regolare e le loro tendenze possono, per molti aspetti, essere anticipate. Questo di per sè potrebbe apparire poco interessante se non fosse che – e questo è forse l'aspetto centrale di tutto questo discorso – proprio la prevedibilità dei processi stocastici ci permette di sfruttarli come punto di riferimento nei processi di misurazione spaziale.

Il tema della distribuzione di Poisson sarà ripreso nei prossimi capitoli.

5
La densità

Una delle misure più utili alla comprensione e allo studio dei peasaggi umani è la densità. Questa, che costituisce una misura di sintesi abbastanza elementare, può essere intesa come un indice di frequenza di un determinato tipo di oggetto o fenomeno all'interno di un'area o intervallo spaziale. Solitamente la densità non viene considerata un indicatore di struttura spaziale (vedi capitolo 6[1]; anche se, in realtà, questo, a seconda del contesto applicativo, potrebbe non essere del tutto vero. Il fatto è che la densità è ritenuta un tipo d'analisi elementare, inadatto a fornire nozioni utili al processo di ricerca e comprensione sullo spazio. Molti la considerano come una semplice misura descrittiva troppo sommaria che di fatto poco può aggiungere. Come verrà illustrato di seguito, questo non è esattamente vero e, di fatto, nella sua semplicità, tale misura può dimostrarsi a volte molto efficace.

La misurazione della densità non fa riferimento solo alla popolazione. Nel campo delle scienze sociali, si può calcolare oltre alla densità delle persone fisiche, anche quella insediativa; ovvero numero di nuclei per unità spaziale o, per esempio, di unità produttive quali fattorie, mulini ecc.

La densità può essere calcolata a partire sia da carte di distribuzione puntuale che da tabelle riassuntive. È fondamentale capire che, in un modo o nell'altro, occorre fare riferimento preciso ad almeno due fattori: la frequenza di un certo fenomeno e il territorio o lo spazio di riferimento: la misurazione della densità richiede un esplicito riferimento ai limiti o ai confini geografici entro i quali questa frequenza sia stata registrata. Di fatto, la densità viene calcolata con:

$$D = \frac{\text{n. elementi}}{\text{area}} = \frac{n}{a} \tag{5.1}$$

Non è infatti sufficiente disporre dei dati relativi alla frequenza del fenomeno studiato, ma è necessario conoscere i confini precisi entro i quali le

[1] Vagaggini 1982, pp. 206-263.

misurazioni sono state realizzate. Questo può rappresentare un problema soprattutto per i processi di raccolta delle osservazioni all'interno di campionature geografiche delle cui caratteristiche non ci sia stata una registrazione dettagliata.

La semplicità di questo computo fa sì che la densità venga utilizzata ampiamente da diversi ricercatori impegnati in qualche ambito dell'analisi e interpretazione delle carte di distribuzione. Accade perciò che in alcune discipline storiche ci si limiti spesso a fornire la densità quale unica indicazione o caratteristica distintiva di un determinato sistema di stanziamento. Se fornita senza maggiori indicazioni, la densità non rappresenta né un dato significativo, né tanto meno un indicatore concreto delle caratteristiche distintive delle reti di stanziamento.

Come già detto in precedenza, la densità non è però un indicatore di struttura. Ovvero, essa non fornisce, né può fornire, indicazioni delle caratteristiche distributive delle maglie di stanziamento. La figura 5.1, ad esempio, mostra due distribuzioni spaziali relative ad un determinato fenomeno. Entrambe le distribuzioni sono composte dallo stesso numero di punti o oggetti, ed entrambe le reti si riferiscono a campioni spaziali di eguale estensione. In un contesto simile, il calcolo della densità nei casi A e B fornirà lo stesso risultato. Nonostante ciò è evidente che i due fenomeni presentano delle differenze sostanziali per quanto riguarda la loro struttura spaziale. Ovvero, le loro caratteristiche distributive presentano delle differenze marcate che occorre analizzare in modo da giungere ad una migliore comprensione del fenomeno.

5.1 Calcolo della densità

Ad esempio, nel caso della provincia di Siena si può arrivare a calcolare velocemente la densità degli abitanti per chilometro quadro nel 1951. Partendo dai dati riassunti nella tabella 4.4 si osserva come il totale degli abitanti registrati nel censimento di quell'anno sia di 277.437. Utilizzando come estensione del territorio provinciale 3819,8 chilometri si ottiene:

$$D = \frac{277.437}{3819,8} = 72,6 \text{ abitanti per } km^2$$

Da questo esempio si può concludere come il calcolo della densità, se impiegato solo in questi termini, venga a costituire esclusivamente una misura di tipo riassuntivo. Questo perché, come è stato detto in precedenza, costituisce esclusivamente una misura del rapporto tra frequenza e spazio. Nonostante questo limite, il principale contributo della densità è quello di permettere il processo di confronto tra contesti spaziali eterogenei. È proprio questa sua ca-

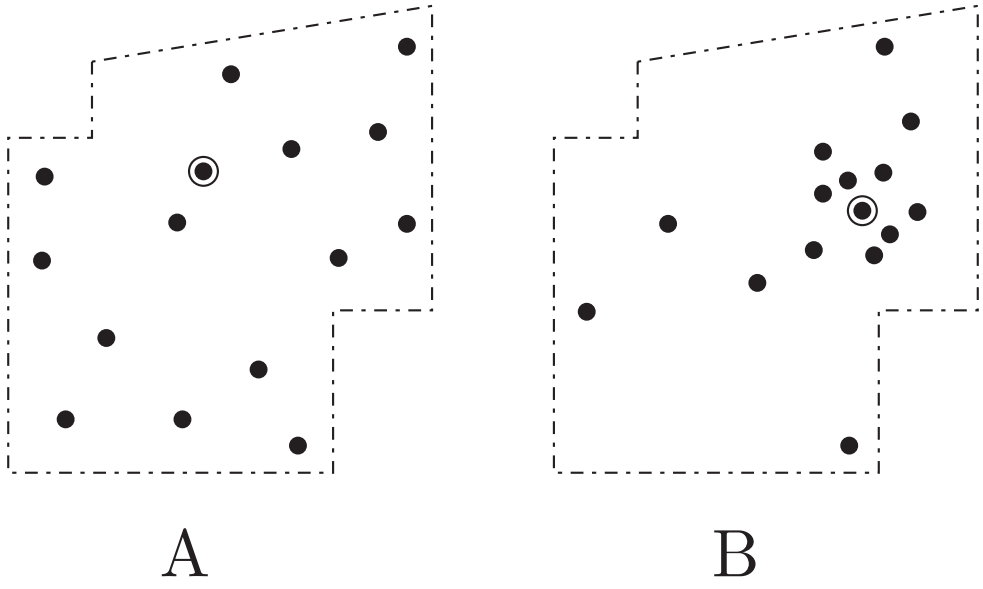


Fig. 5.1. Nei casi A e B le densità sono identiche dato che la superficie dei campioni e la frequenza dei fenomeni sono identiche. Nonostante ciò, la struttura spaziale dei due esempi è evidentemente diversa. Di per sè la densità intesa come valore assoluto – al di fuori di un'articolazione spaziale più ampia – rappresenta una misura sostanzialmente generica.

ratteristica che rende questo tipo di misurazione essenziale per qualsiasi studio sul paesaggio umano.

Ed è in questo senso che si richiederà al lettore di realizzare un collegamento con quanto suggerito nel capitolo 2 relativo ai modelli generali e alla loro utilità nella costruzione di nuova conoscenza. Due unità territoriali o campioni spaziali differenti a livello di forma ed estensione possono essere comparati in base alla densità di un determinato fenomeno rilevato al loro interno. Così la densità della popolazione di un'intera regione può essere confrontata con quella di una provincia o perfino con quella di un comune. In altre parole, dato che si tratta di valori di frequenza relativi, essi possono essere sfruttati per confrontare due entità territoriali diverse o in tempi diversi.

Per capire meglio questo aspetto si osservi la figura 5.2. In questa sono rappresentati due campioni (C e D) molto differenti tra loro. Il primo campione è diverse volte più piccolo del secondo. Al di là della forma schematica dei due esempi, la geografia umana, la storia degli insediamenti e l'archeologia dei paesaggi si trovano frequentemente di fronte a condizioni di questo tipo.

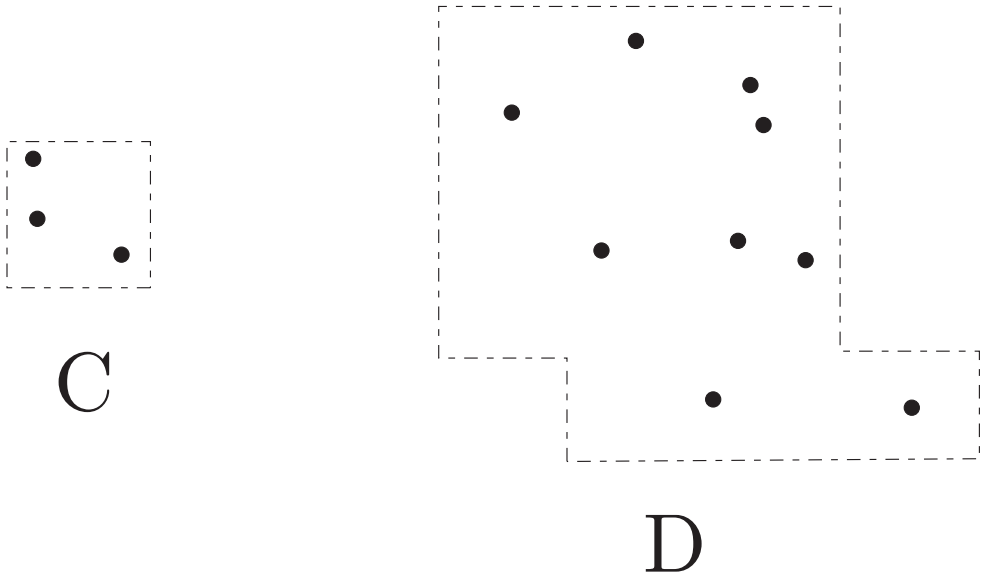


Fig. 5.2. Grazie al calcolo della densità è possibile confrontare formalmente entità spaziali differenti sia per forma che per struttura. Di fatto, il numero di abitanti di un campione non corrisponde alla sua frequenza in termini spaziali. In questo caso, ad esempio, anche se i campioni C e D sono molto diversi, è possibile stabilire che la densità del primo è superiore a quella del secondo.

La figura illustra i siti o le frequenze di insediamenti rilevati durante una ricerca al loro interno. Osservando il grafico si possono capire anche le difficoltà pratiche riscontrabili nel raffronto delle due realtà. Arbitrariamente si potrebbe affermare che il campione D registra un numero maggiore di insediamenti al suo interno. Vale lo stesso per quanto riguarda la densità? In altri termini, è corretto affermare che gli insediamenti sono più frequenti nel campione D che nel C?

Dal momento che si possiedono sia i dati relativi alla frequenza dei siti che all'estensione dei campioni, è possibile calcolare la densità.

$$D_C = \frac{3}{4,3km^2} = 0,69 \text{ siti per } km^2$$

mentre

$$D_D = \frac{9}{35,2km^2} = 0,25 \text{ siti per } km^2$$

Come si può osservare, la densità (o frequenza relativa) del campione C è quasi tre volte quella del campione D. Ci si trova dunque di fronte ad una

			Densità		
n	comune	area km^2	1951	1961	1971
1	Abbadia San Salvatore	59.1	119.3	144.6	144.1
2	Asciano	215.8	45.1	37.4	27.2
3	Buonconvento	64.5	68.9	58.5	46.8
4	Casole d'Elsa	148.4	35.5	28.1	20.4
5	Castellina in Chianti	99.8	49.0	36.6	29.2
6	Castelnuovo Berardenga	177.0	56.1	44.3	28.9
7	Castiglione d'Orcia	141.5	37.1	32.9	25.2
8	Cetona	53.1	91.6	80.8	64.0
9	Chianciano Terme	36.7	123.9	149.6	184.9
10	Chiusdino	141.9	34.0	28.6	20.2
11	Chiusi	58.1	149.3	152.3	151.0
12	Colle di Val d'Elsa	92.3	130.7	139.6	160.5
13	Gaiole in Chianti	129.1	42.1	30.8	22.4
14	Montalcino	243.7	41.9	36.2	25.8
15	Montepulciano	165.8	104.7	95.4	86.6
16	Monteriggioni	99.3	52.9	44.2	55.5
17	Monteroni d'Arbia	105.8	52.9	46.6	44.9
18	Monticiano	109.3	27.2	21.4	17.2
19	Murlo	114.6	30.3	22.0	16.5
20	Piancastagnaio	69.6	76.5	80.3	67.5
21	Pienza	122.6	38.9	32.3	24.4
22	Poggibonsi	70.6	203.7	263.8	359.4
23	Radda in Chianti	80.2	36.6	24.3	19.8
24	Radicofani	118.4	23.2	19.3	13.6
25	Radicondoli	133.0	24.3	16.6	9.9
26	Rapolano Terme	82.9	66.8	69.7	62.6
27	San Casciano dei Bagni	92.1	41.5	35.1	26.7
28	San Gimignano	138.3	81.7	72.6	55.5
29	San Giovanni d'Asso	66.4	49.1	35.4	21.4
30	San Quirico d'Orcia	42.2	55.2	54.4	53.6
31	Sarteano	85.1	55.3	50.0	44.2
32	Siena	118.5	443.6	518.7	553.9
33	Sinalunga	78.4	138.2	144.7	143.8
34	Sovicille	143.6	53.4	45.2	37.4
35	Torrita di Siena	58.1	114.9	111.3	116.1
36	Trequanda	63.9	55.0	41.4	25.2
	Totale	3819.8	72.6	70.7	67.3

Tabella 5.1. Densità abitativa dei comuni della provincia di Siena.

situazione completamente distinta da quella che inizialmente, a occhio nudo, si sarebbe potuto intuire. La misurazione ha permesso di osservare la presenza di un'anomalia: se si tratta dello stesso fenomeno, ad esempio di fattorie, ci si chiederà dunque come mai nel campione C si venga a registrare un livello così elevato di frequenza. Nell'ambito di fenomeni di questo tipo, si tratta di aspetti fondamentali che sono emersi solo dopo avere accuratamente identificato le tendenze a seguito del processo di misurazione. In assenza di una procedura formale niente di tutto questo sarebbe emerso.

5.2 L'uso integrato della densità

Come è stato già sottolineato, la densità viene comunemente considerata una misura elementare – dunque troppo sintetica – per potere fornire elementi utili alla comprensione dei fenomeni studiati. Questo giudizio non è però tanto riferibile a questo tipo di quantificazione quanto al tipo di utilizzo che di essa si fa.

Il corretto uso di questa misura può offrire invece uno strumento insostituibile per la comprensione dell'occupazione sociale dello spazio che può andare ben al di là delle possibilità di confronto tra due diversi campioni spaziali, come illustrato nel paragrafo precedente. In altre parole, la densità può essere impiegata per comprendere meglio le caratteristiche distributive dei fenomeni anche in relazione ad un unico territorio sotto esame. Per corretto utilizzo si intende qui l'adozione di una procedura formale e articolata. Questo principio prevede il calcolo della densità sia a livello generale che particolare o locale. È infatti nel confronto tra valori generali e valori particolari che tale misura può fornire il suo più importante contributo.

Si prenda il caso della provincia di Siena. È già stato stabilito che la densità generale di abitanti nel 1951 era di 72,6 abitanti per km^2. Al fine di rendere più efficace l'applicazione della densità occorre procedere al calcolo dei valori di frequenza, sia a livello generale che a livello locale. Questo in modo da potere giungere ad una sorta di modello o paradigma di frequenza generale al quale le singole aree tendono a somigliare o a differenziarsi.

Naturalmente questo tipo d'approccio è possibile lì dove l'estensione generale in relazione alla frequenza della popolazione sia tale di permettere la definizione di aree locali o piccoli campioni con cui essere confrontata.

In una prima fase si calcola semplicemente la densità generale. Ad esempio, il calcolo della densità degli abitanti della provincia di Siena nel 1951.

In una seconda fase si procede invece al calcolo della densità per ogni singolo comune (vedi la tabella 5.1), grazie a estensione e numero di abitanti. Osservando la tabella, si può notare che i valori relativi alla densità tendono

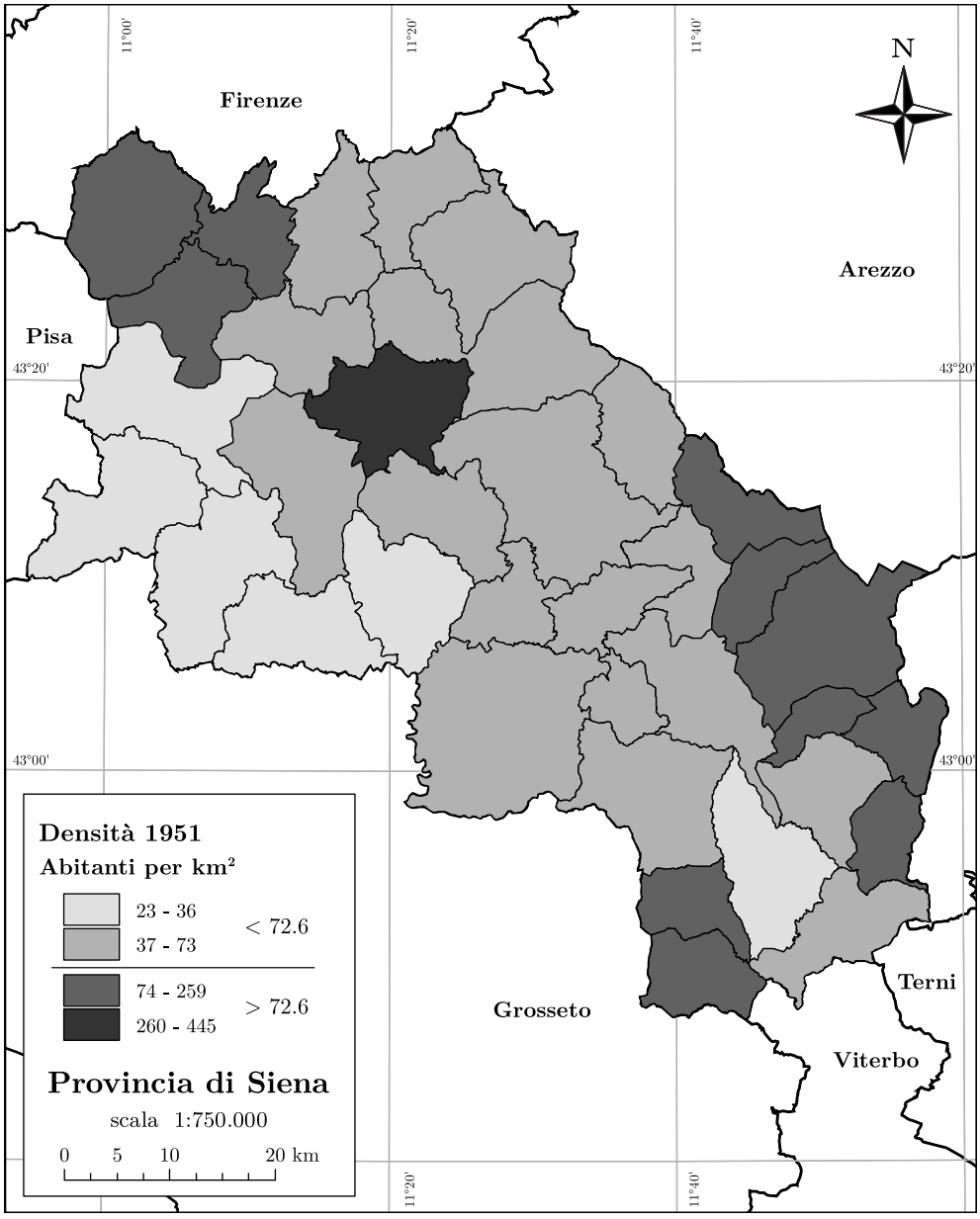


Fig. 5.3. Carta della densità abitativa dei comuni della provincia di Siena nel 1951. Le prime due classi cartografiche sono riferite ai comuni con una densità inferiore a quella generale della provincia; le ultime due, invece, a quelli con valori superiori.

a differenziarsi anche marcatamente. È comunque vero che, pur essendo per alcuni comuni inferiori e per altri superiori, i valori restano, nel caso del 1951, attorno alla densità generale di 72.6 abitanti per chilometro quadro.

In questo caso il valore generale va ad assumere un indice di riferimento. In altri termini, 72.6 viene a costituire un punto che indica le tendenze generali di frequenza degli abitanti nel territorio. I singoli valori per comune invece possono essere assunti come la misura di quanto ciascuno si discosti dalla tendenza generale.

Questo approccio assume un ruolo molto importante soprattutto se i dati della tabella 5.1 vengono presentati a livello cartografico. Nella figura 5.4 si può osservare la classificazione cartografica dei valori di densità. Sono state create quattro classi di cui due con valori al di sopra e due con valori al di sotto della densità generale. Pur nella semplicità di questa operazione, il risultato a livello cartografico permette di cogliere elementi chiari relativi alle differenze anche in termini territoriali. In altre parole, la carta ci permette di rilevare non solo la concentrazione del comune di Siena, ma anche l'aggregazione dei comuni della Val di Chiana e della Val d'Elsa; oppure l'omogeneità della fascia centrale che si muove da nord a sud.

Al di là della lettura ed interpretazione di questa carta, quello che si desidera segnalare è come, arrivati a questo punto, il lettore possa cogliere come diversi elementi quali scala, classificazione e densità, combinate in modo semplice ma formale, abbiano permesso, sulla base di una elementare tabella iniziale (4.1, pag. 143), di ricavare elementi fondamentali; cioè quelli che solitamente si possono chiamare informazioni. Estendendo il dato numerico alla sua dimensione spaziale, si riesce a cogliere un livello interpretativo spesso trascurato.

5.3 Campioni geografici

Un altro modo di calcolare la densità, oltre che integrare valori generali e particolari, è quello di usare metodi discrezionali di suddivisione territoriale. Nel caso appena illustrato, la densità è stata calcolata a partire dai confini comunali.

L'utilizzo di una suddivisione o ripartizione basata su confini politici reali, nel calcolo della densità presenta un inconveniente: la dimensione dell'unità territoriale usata come limite per il computo di tale valore, può condizionare in quei contesti dove lo studio prevede la necessità di lavorare con popolazioni non complete. Mi riferisco in particolare a ricerche nell'ambito della storia degli insediamenti o dell'archeologia dei paesaggi, dove spesso si devono fare i conti con i dati parziali. In questo senso, la natura frammentaria di un campione

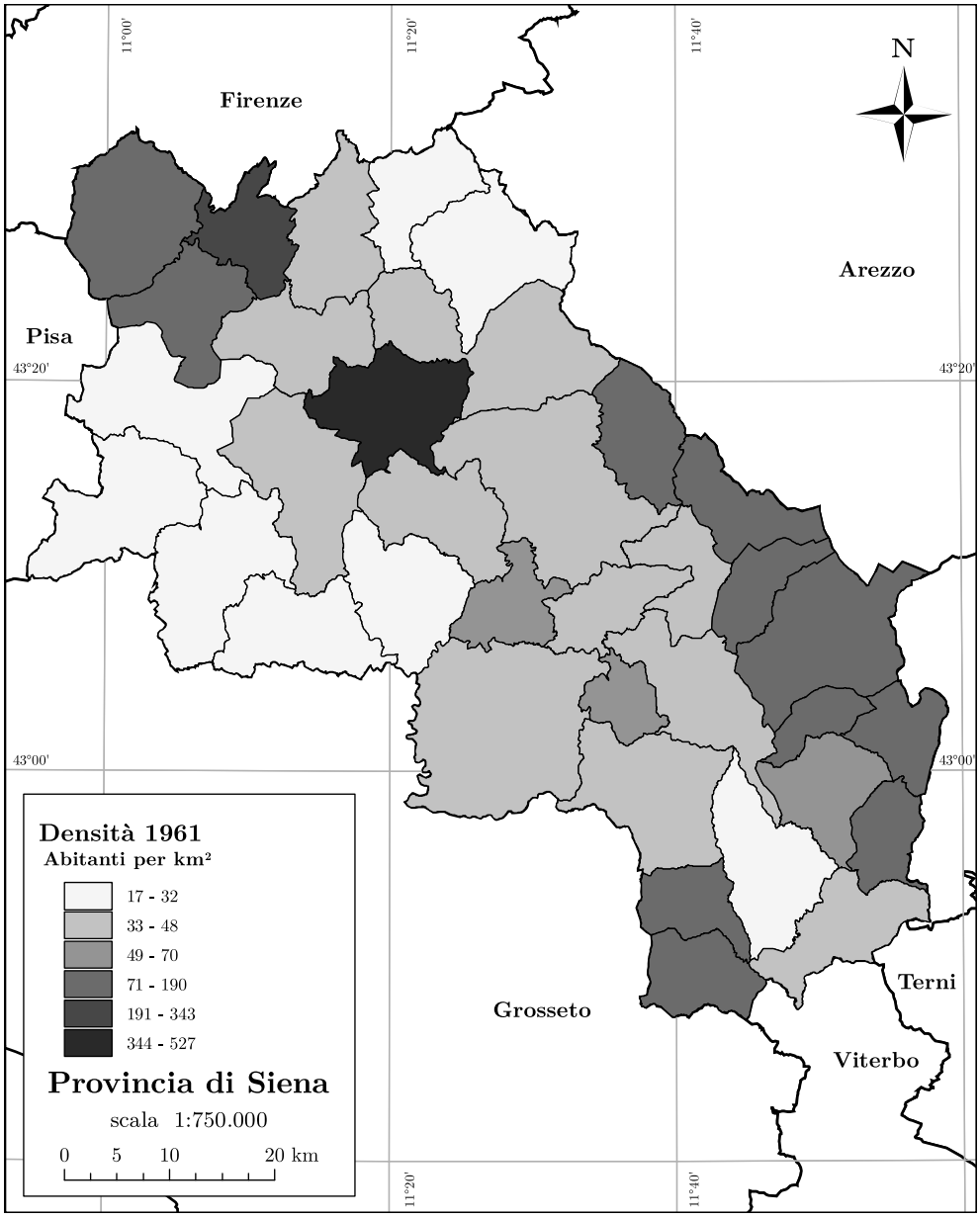


Fig. 5.4. Carta della densità abitativa dei comuni della provincia di Siena nel 1961. Nella carta i valori di densità locale (a livello comunale) sono stati ordinati in sei classi in modo da definirne tre al di sopra e tre al di sotto della densità generale della provincia.

di dati da analizzare potrebbe andare a costituire un problema per quelle porzioni di territorio troppo piccole. In altri termini, la variabilità statistica dei comuni piccoli sarà maggiore che quella dei grandi. Il fatto che si lavori con unità spaziali di diverse dimensioni fa sì che si operi su elementi con grado di affidabilità statistica non omogenea. Così, un comune piccolo che per fattori accidentali veda la comparsa di un villaggio in più all'interno del suo territorio, presenterà un innalzamento eccessivo del valore di densità. Nello stesso modo, comuni piccoli rischiano spesso di presentare alla fine delle analisi valori molto bassi o non significativi qualora, per diverse ragioni, non fosse giunta fino al ricercatore la notizia dell'esistenza di un determinato insediamento. Oltre a questo occorre considerare che è facile che, per unità territoriali piccole, la presenza di siti al di qua o là di là dei loro confini possa, in un senso o nell'altro, alterare definitivamente i risultati.

Per questo motivo, in certe condizioni si preferisce utilizzare una suddivisione arbitraria, dove ogni elemento cartografico utilizzato come limite di un territorio possieda aree della stessa estensione. Il metodo più comune è la griglia a quadrati, sulla base della quale ogni frazione territoriale è composta da un quadrato di pari superficie 5.5. La creazione della griglia a quadrati prevede l'utilizzo di una serie di regole che verranno trattate nel prossimo capitolo (paragrafo 6.4).

Con la suddivisione arbitraria si può arrivare perciò ad un dato analitico molto più preciso, visto che si raggiunge una differenziazione dei valori di densità dell'area di studio, e si può procedere da un livello generale a uno particolare. Questo non significa che i dati a livello generale risultino meno importanti o meno affidabili di quelli particolari. È appunto nell'integrazione tra risultati generali e risultati particolari che le misure di frequenza relativa possono raggiungere il maggior grado di formalità statistica nello studio delle distribuzioni spaziali.

Avere la possibilità di calcolare le differenze tra valori attesi e valori osservati è molto importante, visto che esistono alcuni metodi statistici (tra i quali quello del chi-quadro è forse il più importante) che permettono di calcolare se questa differenza sia tale da permettere di ipotizzare delle alterazioni nella struttura spaziale della rete di distribuzione puntuale all'interno dell'area di studio.

5.4 Il test del chi–quadrato (χ^2)

Il calcolo del χ^2 è indubbiamente uno dei "test di significatività" più importanti e più utili allo studioso interessato all'applicazione dei metodi quan-

titativi alla storia dell'insediamento[2]. Un test di significatività può, in parole semplici, essere descritto come una procedura di quantificazione del grado di significatività relativo a certe osservazioni a livello statistico. In altri termini, la misurazione di quanto peso certe affermazioni o valutazioni di tipo statistico possono avere. Normalmente un test statistico si fonda sull'"ipotesi nulla" o H_0. In campo statistico l'ipotesi nulla rappresenta un'affermazione sul mondo reale che può in qualche misura essere confutata. Se l'ipotesi nulla viene confutata dal test di significatività, allora viene accettata l'"ipotesi alternativa" (H_1).

Il χ^2 serve sostanzialmente a identificare un indice di significatività nelle discrepanze tra valori osservati e valori attesi. Il calcolo del χ^2 corrisponde alla procedura per l'individuazione della soglia di confidenza secondo la quale un'ipotesi nulla può essere confutata.

La distribuzione χ^2 rappresenta uno degli strumenti teorici più importanti proprio perché può facilitare il compito di analisi della significatività delle osservazioni. L'elemento che bisogna ricordare è che in molte circostanze il computo delle frequenze o quantità delle osservazioni dimostra di somigliare alla distribuzione χ^2. Questa somiglianza viene sfruttata come punto di riferimento per identificare le soglie di significatività di cui si parlava sopra.

Per spiegare meglio la natura di questo test, si prenda in considerazione il seguente esempio. Si immagini un intervento di scavo archeologico su due aree funerarie: A e B. Dopo l'intervento di scavo e l'analisi in laboratorio, si è stabilito che, degli scheletri estratti dall'area A, 35 sono riferibili a maschi e 27 a femmine. Dall'area B sono stati invece estratti 48 scheletri di maschi e 21 di femmine.

Si potrebbe così formulare l'ipotesi nulla secondo la quale: *la frequenza di maschi e femmine nelle aree funerarie A e B è equivalente.* Questa affermazione potrebbe essere rappresentata da:

$$H_0 : P_m = P_f$$

dove H_0 rappresenta l'ipotesi nulla, P_m la frequenza attesa per scheletri appartenenti ai maschi e P_f quella delle femmine.

L'ipotesi nulla si baserebbe in sostanza sul principio che la popolazione delle aree funerarie sarebbe un riflesso della popolazione reale, dove le possibilità che una persona sia maschio sono pari a quelle che sia femmina. Scopo dell'applicazione del χ^2 non è quello di identificare possibili anomalie nella distribuzione, oppure di stabilire se i dati osservati siano plausibili uscite da un sistema non influenzato, quanto giungere o non giungere alla confutazione

[2] Una introduzione al uso del χ^2 in campo geografico si trova in Mathews 1985, pp. 135-142.

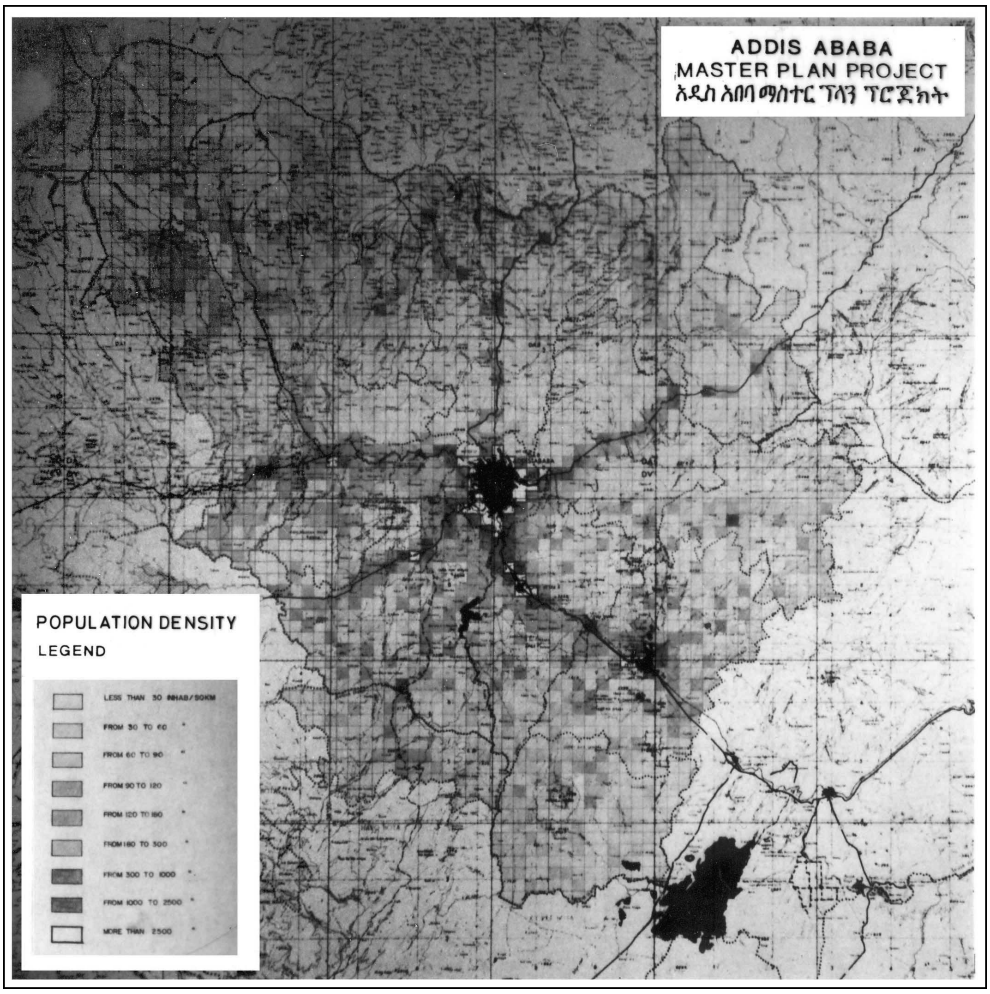


Fig. 5.5. L'utilizzo della suddivisione discrezionale dello spazio costituisce una delle tecniche più efficaci per la misurazione della densità. In questo caso si può osservare la rappresentazione cartografica della densità della popolazione attorno alla città di Addis Abeba. (Greppi manoscritto)

(in termini probabilistici) dell'ipotesi nulla. In altri termini, verificare se le discrepanze tra dati osservati e attesi siano tali da permettere di considerare l'ipotesi nulla non plausibile.

Il calcolo del χ^2 presenta una procedura molto semplice. Infatti esso viene calcolato con:

$$\chi^2 = \sum_{j=1}^{n} \frac{(O_j - E_j)^2}{E_j} \tag{5.2}$$

dove O_j rappresenta i valori osservati ed E_j i valori attesi; ovvero:

$$\chi^2 = \text{sommatoria di} \frac{(\text{frequenze osservate} - \text{frequenze attese})^2}{\text{frequenze attese}}$$

Da questa formula si può per prima cosa notare che il χ^2 ha come funzione quella di accumulare progressivamente una misura di differenze tra i valori osservati e quelli attesi. Il χ^2 alla fine non è altro che un totale di tali differenze. In parole semplici, se i valori osservati saranno vicini a quelli attesi, allora il χ^2 sarà tendenzialmente piccolo; se invece le differenze tra osservazioni e attese sarà grande allora anche il χ^2 presenterà valori elevati.

Dato che il significato di una formula risulta sempre criptico, si osservi subito l'applicazione al problema proposto in precedenza. Prima va sottolineato ancora una volta come lo scopo dell'intera operazione vada nella logica dell'identificazione di differenze significative tra valori osservati e attesi.

Partendo dal caso della necropoli A, il primo passo da compiere è quello di organizzare i dati in una tabella (vedi tabella 5.2), mettendo nella prima colonna i valori osservati (O) e nella seconda quelli attesi (E). La lettera E deriva dal termine *expected*, ovvero "atteso" in lingua inglese. I valori attesi vanno calcolati partendo dalle percentuali possibili di nascita per sesso per la specie umana: 50% per i maschi e 50% per le femmine. Così, per un'area con una presenza di 62 scheletri, ci sarà un'attesa di 31 maschi e 31 femmine. Naturalmente, come si può osservare dai dati raccolti, il numero di maschi e femmine non corrisponde ai valori attesi (E). Col test del χ^2 si intende dunque verificare se tali differenze permettano di confutare l'ipotesi che tali frequenze rispondano ad un modello diverso da quello che vede una distribuzione di 50%-50%.

Dunque, applicando l'equazione appena descritta (vedi 5.2), il valore ottenuto per il χ^2 corrisponde a 1.03:

$$\chi^2 = \frac{(27-31)^2}{31} + \frac{(35-31)^2}{31} = 1.03$$

Area A		
	osservati (O)	attesi (E)
maschi	27	31
femmine	35	31
totale	62	62

Tabella 5.2. Valori osservati (O) e attesi (E) relativi al sesso degli scheletri rinvenuti nell'area A.

Il χ^2 va successivamente confrontato con una tabella di significatività (vedi 5.4). La prima colonna (*d.f.*[3]) presenta i cosiddetti gradi di libertà. Essi indicano, come il nome stesso suggerisce, le possibili "uscite" o "eventi" che una variabile può assume nel processo di classificazione. I gradi di libertà vengono calcolati semplicemente con:

$$d.f. = \text{classi possibili} - 1$$

Ad esempio, nel caso di maschi e femmine, i gradi di libertà saranno calcolati nel seguente modo:

$$d.f. = 2 - 1$$

dato che sono solo 2 le possibili classi. Dunque, nel caso di un cimitero dove la popolazione si divida in due gruppi (maschi e femmine), il numero di gradi di libertà è pari a 1.

Se il valore misurato per χ^2 non supera i valori di significatività vuol dire che le discrepanze studiate tra valori osservati e attesi non sono sufficienti a rifiutare o confutare l'ipotesi nulla H_0; nel caso qui trattato, dell'equivalenza in termini percentuali tra maschi e femmine.

Nell'applicazione del test del χ^2 è necessario stabilire preventivamente con che livello di confidenza si desideri confutare un'ipotesi nulla. In questo caso, si procederà ad adottare una soglia del 95%. Ovvero, si andrà a confutare l'ipotesi nulla qualora le discrepanze tra valori osservati ed attesi presentino un χ^2 superiore a tale soglia.

Così la tabella di significatività indica che il grado di significatività al 95% (terza colonna: 0,05) con grado di libertà 1 è di 3,84. Visto che il χ^2 dell'area A (1,03) non rifiuta l'ipotesi nulla, si può perciò affermare, con un 95% di certezza statistica, che le popolazioni osservate in quest'area funeraria non

[3] Dall'inglese *Degrees of Freedom* ovvero gradi di libertà.

Area B		
	osservati (O)	attesi (E)
maschi	48	34.5
femmine	21	34.5
totale	69	69

Tabella 5.3. Valori osservati (O) e attesi (E) relativi al sesso degli scheletri rinvenuti nell'area B.

confutano l'ipotesi di equilibrio tra maschi e femmine. Fare attenzione al fatto che il χ^2 non permette di affermare che è stato dimostrato l'equilibro, bensì che l'assenza di equilibro tra maschi e femmine non può essere confutata.

Il confronto tra il χ^2 e i valori di siginificatività riportati nella tabella rappresenta il processo di misurazione delle differenze tra valori osservati e valori attesi. In altri termini, se le misurazioni sulla base dei risultati e sulla base delle soglie appaiono molto diversi, allora questa discrepanza può essere utilizzata per confutare l'ipotesi nulla.

Si prenda in considerazione adesso il caso B. Nella seconda area funeraria, il valore individuato con il test del χ^2 è stato di 10,56:

$$\chi^2 = \frac{(48 - 34.5)^2}{34.5} + \frac{(21 - 34.5)^2}{34.5} = 10.56$$

In questo caso, il valore del χ^2 supera abbondantemente il limite di significatività al 95%, di 3,84 con 1 grado di libertà. In questo caso il risultato riscontrato indica che la distribuzione degli scheletri nelle due categorie mostra un'anomalia; ovvero, l'ipotesi di popolazioni equivalenti tra maschi e femmine può essere rifiutata con un grado di attendibilità del 95%.

A questo punto, dai risultati ottenuti si possono costruire delle ipotesi del perché i dati raccolti all'interno dell'area B presentino delle anomalie. Era forse la società studiata un gruppo che prevedeva aree funebri separate a seconda del sesso? Si trattava di un'area funebre dedicata esclusivamente a ceti militari, poi allargata a tutta la popolazione?

5.5 Il χ^2 e la geografia quantitativa

Il tipo di approccio alla base del calcolo del χ^2 – ovvero il confronto tra valori attesi e valori osservati – può essere impiegato per un utilizzo ancora più articolato della densità geografica.

	Soglie critiche del χ^2				
d.f.	**0.1**	**0.05**	**0.025**	**0.01**	**0.005**
1	2.706	3.841	5.024	6.635	7.879
2	4.605	5.991	7.378	9.210	10.597
3	6.251	7.815	9.348	11.345	12.838
4	7.779	9.488	11.143	13.277	14.860
5	9.236	11.070	12.833	15.086	16.750
6	10.645	12.592	14.449	16.812	18.548
7	12.017	14.067	16.013	18.475	20.278
8	13.362	15.507	17.535	20.090	21.955
9	14.684	16.919	19.023	21.666	23.589
10	15.987	18.307	20.483	23.209	25.188
11	17.275	19.675	21.920	24.725	26.757
12	18.549	21.026	23.337	26.217	28300
13	19.812	22.362	24.736	27.688	29.819
14	21.064	23.685	26.119	29.141	31.319
15	22.307	24.996	27.488	30.578	32.801
16	23.542	26.296	28.845	32.000	34.267
17	24.769	27.587	30.191	33.409	35.718
18	25.989	28.869	31.526	34.805	37.156
19	27.204	30.144	32.852	36.191	38.582
20	28.412	31.410	34.170	37.566	39.997
21	29.615	32.671	35.479	38.932	41.401
22	30.813	33.924	36.781	40.289	42.796
23	32.007	35.172	38.076	41.638	44.181
24	33.196	36.415	39.364	42.980	45.559
25	34.382	37.652	40.646	44.314	46.928

Tabella 5.4. Nella tabella sono illustrati i valori critici del χ^2. La prima colonna (*d.f.*) rappresenta i gradi di libertà.

È stato già indicato come questo test possa essere esteso invece a più classi o categorie. Nel caso dell'applicazione allo studio delle maglie di stanziamento, ad esempio, si può ricorrere a classi di tipo spaziale. Infatti molte forme di suddivisione spaziale possono essere sfruttate in modo concreto per l'applicazione del χ^2. In sostanza l'applicabilità si fonda sul principio che i valori di frequenza osservati all'interno di ogni singola suddivisione spaziale possono essere confrontati con quelli attesi.

L'esempio qui illustrato si basa su una suddivisione in *buffers* concentrici. Si segnala che questa è solo una delle possibili forme di suddivisione con le quali si può operare. Di fatto le operazioni si fondano principalmente sull'ipotesi nulla di partenza.

Il caso d'applicazione si basa su ricerche che facevano parte di una serie di studi più ampi condotti al fine di identificare le caratteristiche distributive della maglia dei castelli medievali tra XII e XIV secolo[4]. In questo caso impiegato qui come esempio si partiva dalla seguente ipotesi nulla e ipotesi alternativa:

$$H_0 : \text{ la città non ha effetti sulla struttura della maglia castrense}$$

$$H_1 : \text{ la città ha effetti sulla struttura della maglia castrense}$$

L'intero processo è caratterizzato in primo luogo dalla definizione delle fasce spaziali attorno ai centri urbani. Vale a dire, il frazionamento della geografia in più insiemi spaziali dove compiere le quantificazioni delle densità relative. Il processo di suddivisione geografica naturalmente avrà origine dai nuclei urbani e sarà caratterizzato dalla definizione di fasce circolari concentriche o *buffers*, aventi come centro appunto la città. Queste fasce concentriche sono state definite secondo intervalli regolari di $5km$.

Poi si è proceduto a fondere topologicamente gli anelli in categorie spaziali; ovvero tutti gli anelli tra 0-5 km venivano a costituire un'unica entità spaziale. Questo in modo da potere misurare non la densità attorno ad un'unica città, quanto piuttosto la densità attorno alle città in generale. La procedura è stata ripetuta per tutte le classi in questione: 5-10 km, 10-15 km ... e così di seguito.

Con l'aiuto del GIS e sulla base della distribuzione dei castelli è stato poi possibile calcolare le frequenze relative e confrontarle con quelle attese. La procedura è sintetizzata nella tabella 5.5. La prima colonna (*buffer*) contiene la sequenza delle classi, espressa come anelli (con raggi in km) attorno alle città della regione. La seconda colonna indica l'area in km^2 che ogni singola categoria di anello occupa. Questo valore è molto importante dato che fornisce la proporzione con la quale i valori attesi verranno poi calcolati. La terza colonna presenta il numero di castelli osservati (o) nella classe spaziale o insieme di *buffers* corrispettivi. La quarta il numero di castelli attesi (e) per quella fascia. La quinta colonna illustra la densità osservata (o/a) nella fascia e infine l'ultima colonna quella relativa al calcolo del χ^2.

La procedura si basa sul principio della distribuzione casuale indipendente discussa precedentemente (vedi paragrafo 4.9.3). Secondo i principi di una distribuzione di Poisson, il numero di punti attesi e all'interno di una certa classe di anello dipende esclusivamente dalla sua estensione o superficie. In altre parole rappresenta una proporzione tra l'area totale degli anelli e il totale di punti osservati. Ad esempio per calcolare il numero di insediamenti attesi nel secondo anello:

[4] Macchi 2007.

$$\begin{aligned} 2461.6 &\longrightarrow x \\ 21148.1 &\longrightarrow 903 \end{aligned}$$

ovvero

$$x = \frac{2461.6 \times 903}{21148} = 105.1$$

Alla fine tutti i valori attesi e dovranno corrispondere al totale osservato; in questo caso 903. L'ultima colonna presenta l'operazione $(o-e)^2/e$; infatti sommando tutti i valori di questa colonna si ottiene il valore del χ^2. Nel caso di questa analisi si registra un χ^2 con 7 $d.f.$ dato che sono otto le classi con le quali si è proceduto alla misurazione.

In questo caso specifico le analisi condotte hanno permesso di confutare l'ipotesi nulla illustrata sopra.

		Siti			
buffer	Area (km^2)	Osservati (o)	Attesi (e)	o/a	$(o-e)^2/e$
5	882.9	13	37.7	0.015	16.2
10	2461.6	75	105.1	0.030	8.6
15	3685.4	181	157.4	0.049	3.5
20	4572.2	213	195.2	0.047	1.6
25	4351.9	208	185.8	0.048	2.6
30	2935.1	129	125.3	0.044	0.1
35	1518.4	53	64.8	0.035	2.2
40	740.5	31	31.6	0.042	0.0
Totale	21148.1	903	903.0		χ^2 35

Tabella 5.5. anno 1275: frequenza, densità castrense e χ^2 per *buffers* concentrici attorno ai centri urbani.

gradi di libertà	$\chi^2_{.995}$	$\chi^2_{.99}$	$\chi^2_{.975}$	$\chi^2_{.95}$
7	20,3	18,5	16,0	14,1

Tabella 5.6. Tabella delle percentuali relative ai livelli di significatività del χ^2 con 7 "gradi di libertà".

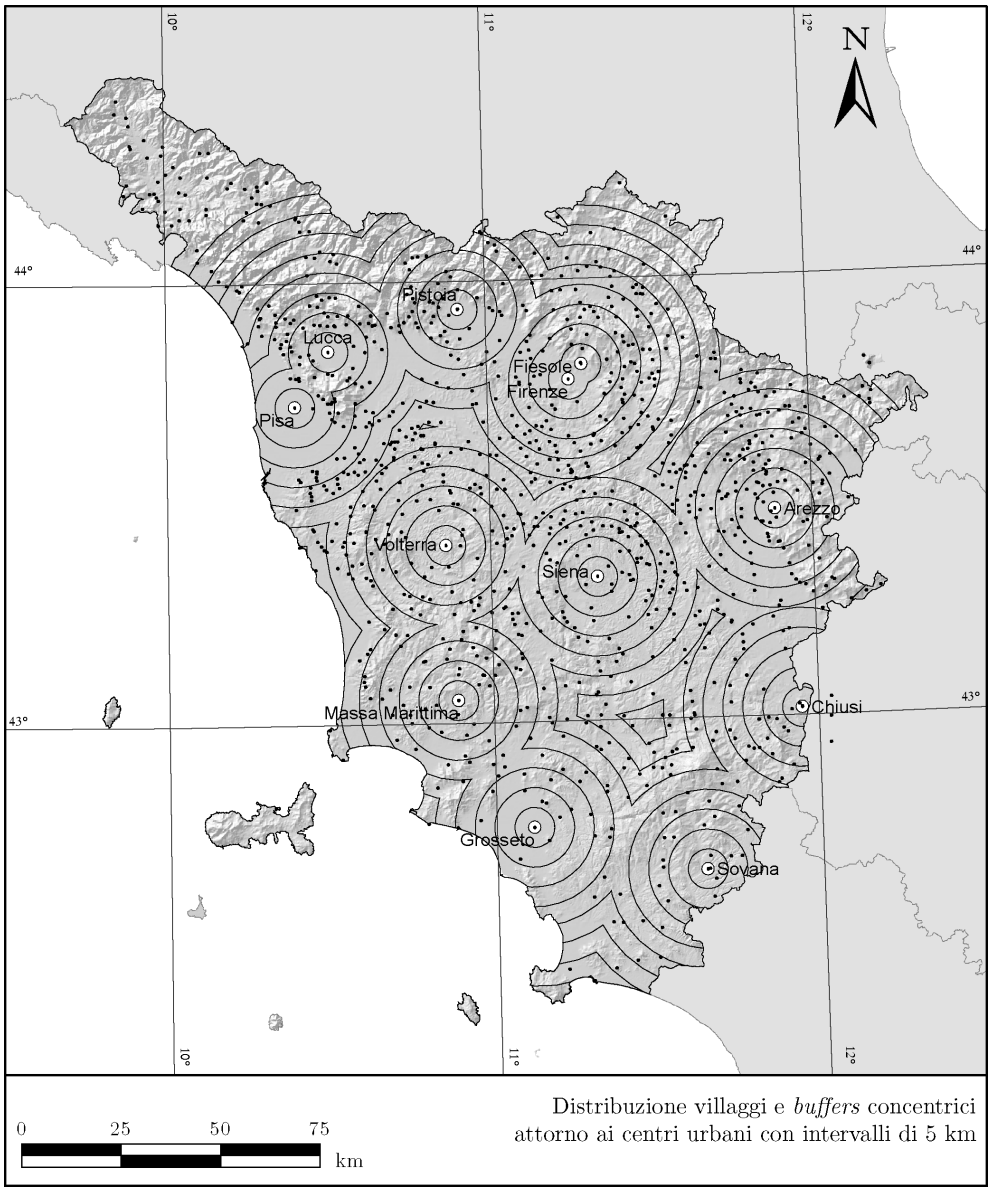


Fig. 5.6. Suddivisione arbitraria della geografia della regione con una serie di anelli concentrici che verranno sfruttati successivamente per la misurazione della densità.

6
Point pattern analysis

Scopo del *point pattern analysis* è quello di identificare le caratteristiche strutturali di una distribuzione per punti. Rappresenta un settore specifico dell'analisi spaziale e comprende più metodi che, a partire dagli anni '50 del secolo scorso, sono stati sviluppati prevalentemente in seno a discipline come la biologia o l'ecologia; in seguito, progressivamente, la geografia li prese prima in prestito e poi li accolse definitivamente nel suo corpo metodologico durante la decade del '60[1]. Si tratta dunque di un settore metodologico trasversale a discipline marcatamente impegnate nella soluzione di problemi direttamente connessi con fattori quali la posizione e la distribuzione. Come verrà illustrato lungo le pagine di questo capitolo, il *point pattern analysis* ha continuato a svilupparsi fino ai nostri giorni cercando di individuare metodi sempre più corretti in termini di misurazione, cercando di espandere le possibilità di compiere confronti più rigorosi e precisi[2].

Tanto è importante il ruolo e peso di questo settore che spesso *point pattern analysis* e analisi spaziale vengono scambiati o considerati equivalenti. Ma, come è stato anticipato sopra, non è così. Oltre all'analisi della distribuzione per punti, l'analisi spaziale si occupa dello studio delle caratteristiche strutturali delle aree e delle reti[3].

Il *point pattern analysis* rappresenta un settore metodologico essenziale per qualsiasi disciplina sociale o storica connessa con lo studio degli insediamenti[4]. Il motivo principale è che gran parte delle ricerche in questo ambito ha a che

[1] Gatrell, Bailey, Diggle, Rowlingson 1996, p. 256; Diggle 1983.

[2] Un resoconto critico dell'evoluzione storica dell'analisi spaziale – e dunque di alcuni aspetti relativi al *point pattern analysis* – si trova in Longley 2000.

[3] Per quanto riguarda l'analisi delle aree, fare riferimento a Unwin 1986, pp. 179-225. L'analisi delle reti ha beneficiato dalla comparsa del compendio *Network analysis in geography*: Chorley, Haggett 1972.

[4] Hammond, McCullagh 1975, p. 237: «Essentially geography is concerned with distributions in space and one of the most important distributions the geographer has to consider is that of human settlement.»

fare con carte di distribuzione per punti e, come è stato suggerito nel primo capitolo, la loro lettura e comprensione risultano compiti impossibili senza l'adozione di un adeguato approccio formale. Nei prossimi paragrafi verranno dunque introdotti i principali elementi metodologici di questo approccio.

6.1 Le carte per punti

Solitamente lo studioso si avvale di carte per punti per rappresentare la distribuzione di oggetti e fenomeni nello spazio. L'utilizzo della carta con simboli puntuali rappresenta una soluzione che prevede in ogni caso un notevole grado d'astrazione[5]. Questo perché si procede a rappresentare oggetti reali, come edifici, fattorie, chiese o villaggi con simboli cartografici puramente astratti[6]. Geometricamente, infatti, un punto non possiede nessuna dimensione. Viceversa, il punto può essere sfruttato per rappresentare una posizione nello spazio[7]. Posizione che può essere indicata con un elevato grado di precisione grazie alle posizioni decimali delle coordinate del punto.

La precisione non costituisce però una qualità assoluta nel campo dell'analisi delle carte per punti. Infatti il grado di precisione dipende in larga misura dalla scala cartografica sulla quale si vanno ad operare le quantificazioni[8]. Si può affermare che nel contesto di un unico processo di analisi, più che l'utopistica ambizione ad una precisione assoluta, possa contribuire al raggiungimento di risultati validi una omogeneità degli errori.

Generalmente i ricercatori che all'interno dei propri processi d'indagine fanno uso di carte a punti, sono fortemente interessati a problemi di posizione o distribuzione all'interno dello spazio geografico. Lo scopo principale dell'analisi spaziale è l'identificazione delle caratteristiche intrinseche della struttura distributiva di un certo fenomeno[9]. Nel campo dell'analisi spaziale la posizione di ogni elemento di una maglia di stanziamento acquista un ruolo fondamentale. Ed è proprio per questo motivo che l'unico dato indispensabile per portare

[5] Ripley 1981, pp. 3-6.

[6] O'Sullivan, Unwin 2003, pp. 79-80.

[7] Per un'attenta analisi dei dati spaziali e del loro utilizzo nell'ambito dell'analisi spaziale fare riferimento a Fotheringham, Brunsdon, Charlton 2000. pp. 15-22 e pp. 134-138.

[8] Sul tema della scala geografica e della precisione delle misurazioni in campo geografico consultare Mandelbrot 1967, pp. 636-368.

[9] Landini, Massimi 1982, p. 13:«... le strutture matematiche sono in grado di far emergere il messaggio di fondo contenuto in un insieme di particolari dispersi, attributi o essi stessi oggetti spaziali, soltanto se esiste una possibilità reciproca di traduzione fra un certo linguaggio scientifico specialistico [...] che fornisce i modelli interpretativi-descrittivi dei processi soggiacenti, generatori di sequenze dei dati, ed il linguaggio matematico che formalizza il modello e rende le descrizioni logicamente coerenti, confrontabili [...] ed oggettive [...]. »

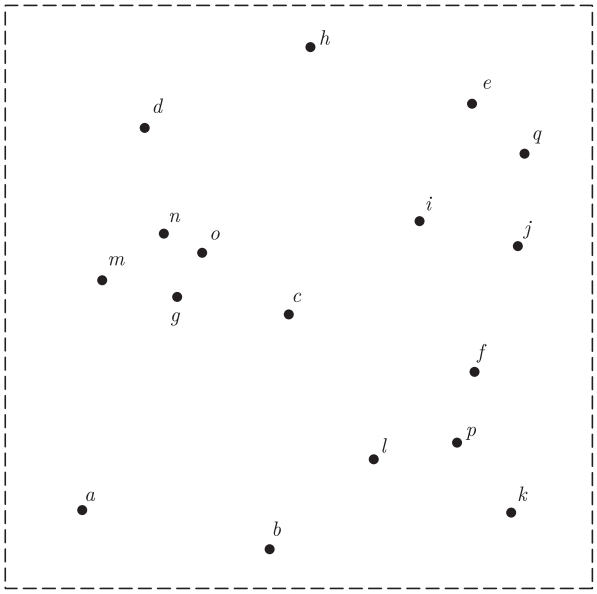


Fig. 6.1. Carta di distribuzione a punti. Le carte di distribuzione per punti rappresentano un modello con un elevato livello d'astrazione: solo gli elementi relativi alla localizzazione o posizione sul piano vengono presi in considerazione. Questo tipo di documento si presta, data la sua semplificazione, ad essere analizzato con i metodi tipici del *point pattern analysis*.

a compimento un processo d'analisi spaziale è una carta di distribuzione per punti[10]; come ad esempio quella illustrata nella figura 6.1.

Le carte di distribuzione per punti non solo rappresentano la distribuzione di oggetti o fenomeni sociali nello spazio; esse possono perfettamente essere utilizzate per rappresentare la posizione di punti di rilevazione relativi all'ambiente e alla geografia fisica. In questo caso, la carta compie un'associazione tra posizione spaziale e misurazione. Un esempio di questo potrebbe essere una carta delle temperature o una delle precipitazioni. Le carte per punti possono essere adeguatamente utilizzate, grazie all'interpolazione cartografica, per modellare misurazioni sul mondo fisico: temperature, velocità del vento, acidità delle precipitazioni[11]. L'aspetto su cui occorre insistere è che nell'analisi spaziale si fa un largo uso delle carte di distribuzione per punti proprio per-

[10] Per una chiara introduzione al tema dell'applicazione della geografia quantitativa allo studio dei processi storici si rimanda a Matthews 1985, pp. 72-76.

[11] Per un'introduzione ai metodi di interpolazione cartografica nell'ambito della geografia quantitativa si rimanda a O'Sullivan, Unwin 2003, pp. 220-241. Inoltre si consiglia la lettura di Fotheringham, Brundson, Charlton 2000, pp. 42-45.

ché l'interesse è fondamentalmente indirizzato allo studio della posizione e dell'organizzazione delle reti di distribuzione o delle misurazioni.

6.1.1 Le carte a punti e la coppia di variabili *x y*

Com'è stato detto in precedenza, l'analisi spaziale è un settore particolare dell'analisi dei dati. Per molti versi risponde ai criteri e alle necessità della statistica descrittiva; solo che, nel caso dell'analisi spaziale, gran parte delle operazioni sono compiute ad un livello caratterizzato da due dimensioni invece che da una. Di conseguenza, più che la rappresentazione grafico-simbolica (nel senso di cartografica) di una distribuzione a punti, le operazioni d'analisi richiedono la sintesi numerica di tali punti sotto forma di coordinate x e y[12]. In altre parole, i processi d'analisi spaziale, ed in particolare le operazioni che rientrano all'interno del cosiddetto *point pattern analysis*, richiedono numeri da poter essere elaborati statisticamente. Nella quasi totalità i concetti statistici indispensabili per l'applicazione dei metodi per l'analisi delle carte a punti sono già stati introdotti nei capitoli precedenti.

Una rete di distribuzione come quella illustrata nella carta a punti della figura 6.1 può essere sintetizzata con dati numerici. La tabella 6.1 illustra le coordinate dei punti che compongono il sistema di distribuzione sotto forma di coordinate x e y.

6.2 Il baricentro

Diverse misure analitiche sono state sviluppate dalla geografia per descrivere le caratteristiche generali e, in particolare, la tendenza centrale di una rete di distribuzione. Il computo di queste misure di tendenza centrale non presenta particolari difficoltà[13].

Il primo tipo di misurazione che può essere fatto su una rete di distribuzione puntuale è il calcolo del "baricentro" o "centro medio"[14]. Questa misura va intesa come il luogo o la posizione assoluta – vale a dire non corrispondente ad uno dei punti reali della maglia – attorno alla quale una rete di distribuzione per punti tende ad aggregarsi.

Il baricentro è basato sulla media delle coordinate (sia sulle ascisse che sulle ordinate) dei punti. La media delle coordinate di una data distribuzio-

[12] Unwin 1986, pp. 80-81; Fotheringham, Brunsdon, Charlton 2000, pp. 134-138.
[13] O'Sullivan, Unwin 2003, pp. 78-80.
[14] Da *mean center* in lingua inglese.

punti	x	y
a	1.57	1.45
b	4.81	0.76
c	5.15	4.86
d	2.63	8.12
e	8.34	8.53
f	8.39	3.88
g	3.20	5.19
h	5.53	9.52
i	7.45	6.48
j	9.13	6.07
k	9.01	1.38
l	6.63	2.34
m	1.91	5.45
n	2.98	6.26
o	3.62	5.93
p	8.08	2.62
q	9.20	7.67

Tabella 6.1. Coordinate dei punti illustrati nella figura 6.1. Solitamente sono dati come quelli illustrati in questa tabella (e non la loro rappresentazione cartografica) ad essere processati con i metodi del *point pattern analysis*.

ne rappresenta un luogo di particolare interesse nello studio di una rete di distribuzione[15].

Il baricentro di una carta a punti può essere ottenuto semplicemente calcolando la media delle coordinate X e quella delle coordinate Y di un insieme finito di punti. La seguente formula descrive il calcolo del baricentro:

$$(\bar{X}, \bar{Y}) = \left(\frac{\sum_{j=1}^{n} X_j}{n} \; , \; \frac{\sum_{j=1}^{n} Y_j}{n} \right) \tag{6.1}$$

ovvero

$$\text{baricentro} = (\text{media delle X}, \text{media delle Y})$$

Qualora sia necessario identificare il baricentro di una rete di punti con pesi diversi, si dovrà ricorrere al calcolo delle medie ponderate di X e Y. Come è possibile calcolare la media ponderata, è possibile anche ricavare il

[15] Per una dettagliata analisi dei limiti del baricentro si rimanda a Silk 1979, pp. 22-26.

baricentro ponderato. Questo calcolo naturalmente prevede la presenza di una variabile (solitamente indicata come p) contenete i valori associati alla coppia di coordinate[16]:

$$(\bar{X}, \bar{Y}) = \left(\sum_{j=1}^{n} X_j p_j \Big/ \sum_{j=1}^{n} p_j \ , \ \sum_{j=1}^{n} Y_j p_j \Big/ \sum_{j=1}^{n} p_j \right) \tag{6.2}$$

dove X_j rappresenta la variabile relativa alle x delle coordinate della distribuzione, Y_j alle y e p_j una variabile contenente un peso associato ad ogni punto; normalmente, nel campo della geografia umana, la popolazione. Per quanto imponente possa apparire, questa equazione vuol dire semplicemente:

$$\text{baricentro ponderato} = (\text{media ponderata delle } X, \text{media ponderata delle } Y)$$

Il baricentro non rappresenta un indicatore di struttura, bensì una misura di sintesi come d'altronde tutte le misure di posizione. In altre parole, il baricentro non fornisce indicazioni su come la maglia di stanziamento sia organizzata a livello spaziale. Certamente, un inconveniente di questa misura è il fatto che il baricentro è sensibile alle variazioni dei valori estremi così come il calcolo della media di una semplice serie di dati numerici.

Partendo dalle coordinate illustrate nella tabella 6.2 relative ai toponimi indicati come castelli nella *Tavola delle Possessioni*[17], applicando l'equazione del baricentro ponderato (6.3), è possibile ottenere:

$$(\bar{X}, \bar{Y}) = (1699038, 4776374)$$

Nella tabella indicata, la variabile p indica il numero di case registrate nella fonte catastale.

6.2.1 Deviazione della distanza standard

A partire del baricentro è possibile calcolare anche la cosiddetta "deviazione della distanza standard" rappresentata dal simbolo σ_D. Il calcolo di tale misura si ottiene a partire dalla seguente formula:

$$s_D = \sqrt{\left(\sum_{j=1}^{n} d_i^2 \right) \Big/ n} \tag{6.3}$$

[16] Una trattazione dettagliata del metodo si trova in Vagaggini 1982, pp. 243-246.
[17] Per una introduzione alla *Tavola delle Possessioni* consultare Passeri, Neri 1994.

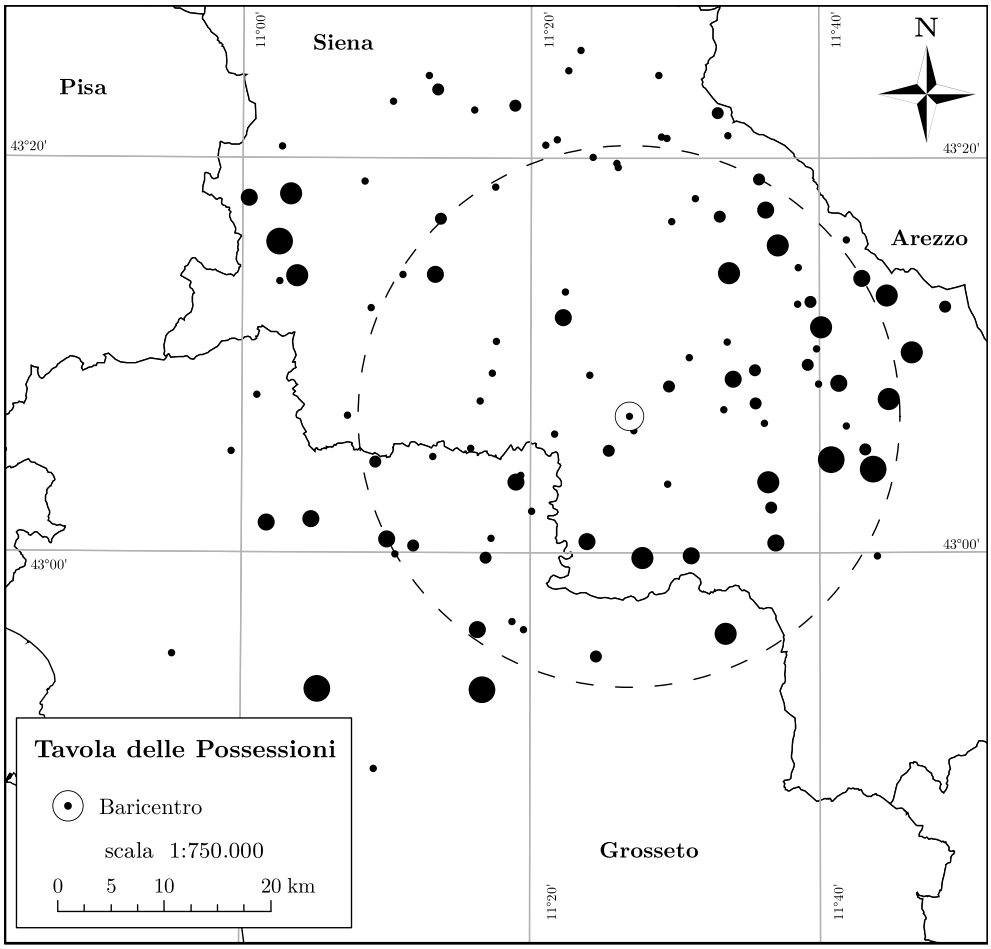


Fig. 6.2. Insieme ai punti relativi ai castelli della *Tavola delle Possessioni* è rappresentato anche il baricentro. Come si può osservare, il baricentro è visibilmente spostato a sud-est rispetto al centro naturale delle maglia dei punti. Il cerchio tratteggiato rappresenta invece la "deviazione standard della distanza". Vedi Olivelli 2006.

dove d rappresenta la distanza di ogni punto della rete analizzata dal baricentro. In altre parole:

$$s_D = \text{scarto quadratico medio delle distanze dei punti dal baricentro}$$

A livello cartografico s_D va rappresentato come un cerchio con un raggio pari alla distanza ottenuta attorno al baricentro. Come qualsiasi misura di dispersione permette di identificare alcuni aspetti relativi alle caratteristiche distributive della maglia. La deviazione della distanza standard va intesa come la rappresentazione cartografica di un indice statistico di dispersione (vedi paragrafo 4.8) dei punti dal proprio baricentro[18].

6.2.2 L'utilizzo del baricentro nella ricerca storica

Il calcolo del baricentro è un tipo di misurazione che non presenta particolari difficoltà a livello computazionale. Non vi è dubbio però che il baricentro costituisca, per ovvi motivi, un indicatore di sintesi modesto. Forse per questo motivo, il baricentro, se confrontato con altri tipi d'analisi, sembra non destare particolare interesse nel campo della geografia quantitativa. Nei testi specializzati e manuali spesso esso viene trattato sbrigativamente per passare a temi apparentemente più importanti. Accade invece che il calcolo del baricentro costituisca una delle misure più importanti per l'analisi e lo studio dell'evoluzione di sequenze storiche delle maglie di stanziamento umano.

L'approccio metodologico consiste semplicemente nella ripetizione del calcolo del baricentro per l'evoluzione storica o la sequenza cronologica di una rete di stanziamento. In questo senso la storia del baricentro e la distribuzione cartografica delle sue diverse posizioni nel tempo permettono di cogliere elementi di grande interesse[19].

La figura 6.3, ad esempio, mostra il percorso evolutivo della tendenza centrale del popolamento negli Stati Uniti d'America a partire dal 1790 al 1930. Come si può osservare dalla carta il baricentro subisce uno spostamento costante, per il secolo e mezzo d'intervallo, a partire dalla costa atlantica fino alla zona dei grandi laghi. Scegliendo ogni punto singolarmente, si può dire ben poco: esso rappresenta una sintesi generale della tendenza. Ma la sua evoluzione storica dà un chiaro segnale delle trasformazioni (anche a livello strutturale) che il popolamento sta subendo nel tempo.

La rappresentazione cartografica dello sviluppo temporale del baricentro, quando riferita a fenomeni reali, come ad esempio una rete d'insediamento o

[18] Unwin 1986, pp. 95-96.

[19] Un saggio rilevante sul tema dell'applicazione dello studio dell'evoluzione storica nel campo della geografia umana è quello di Sviatlovsky, Eells 1937.

Castelli della Tavola delle Possessioni					
X_j	Y_j	p_j	X_j	Y_j	p_j
1665726	4800738	2	1680146	4806533	29
1665853	4788095	1	1728377	4787519	28
1675986	4805247	4	1706920	4795376	63
1663979	4777369	1	1716440	4783174	17
1694923	4800512	11	1715635	4781668	38
1690443	4801568	2	1721284	4773927	39
1685818	4797485	4	1710679	4781047	26
1692357	4808624	1	1707907	4777190	16
1693464	4810519	1	1675976	4742544	22
1661766	4772006	1	1702691	4779303	60
1656723	4752829	1	1685910	4762692	64
1692090	4774412	1	1684998	4777340	19
1683685	4804620	8	1672595	4775631	19
1673542	4797663	8	1675329	4771377	53
1702426	4794688	9	1679088	4763643	26
1697148	4800057	6	1719442	4775993	17
1701788	4802533	1	1690084	4767162	21
1697264	4799635	6	1715677	4787592	71
1707457	4802965	1	1722692	4763935	21
1692669	4787787	2	1710551	4798959	42
1699484	4774955	2	1697214	4773076	62
1718872	4793502	1	1662772	4795809	116
1714436	4790771	8	1666661	4796314	203
1704627	4796919	2	1667472	4788674	226
1714518	4787335	5	1712295	4770534	232
1716730	4779910	9	1722866	4788407	206
1708001	4783590	2	1720461	4789937	154
1711754	4776063	1	1711244	4796085	99
1702814	4770038	8	1712526	4792808	180
1677361	4789011	1	1716789	4785251	191
1688570	4756708	2	1718605	4780021	163
1688989	4770456	2	1725397	4783141	241
1689667	4755979	11	1723369	4778697	219
1686378	4764498	3	1708666	4780092	92
1684202	4772867	3	1713167	4764840	161
1686311	4782989	9	1708669	4756173	197
1686055	4779960	6	1705207	4763392	109
1695199	4780097	10	1700639	4763083	188
1677390	4762729	7	1695441	4764432	154
1680703	4772030	2	1685395	4755878	97
1704467	4782029	1	1688547	4769807	144
1686411	4749944	1	1680414	4789126	88
1679287	4807736	3	1692542	4785396	102
1674518	4785805	15	1669413	4765808	104
1691492	4802103	14	1665228	4765344	130
1687424	4805197	27	1665709	4791860	310
1696576	4753707	31	1722098	4772091	363
1712612	4768172	37	1718134	4772864	324
1706461	4805050	42	1707976	4790049	247
1700833	4808402	17	1670449	4749923	343
1680786	4794416	26	1685959	4750292	367
1701233	4802659	15			

Tabella 6.2. Coordinate Gauss Boaga degli insediamenti citati come castelli nella *Tavola delle Possessioni*.

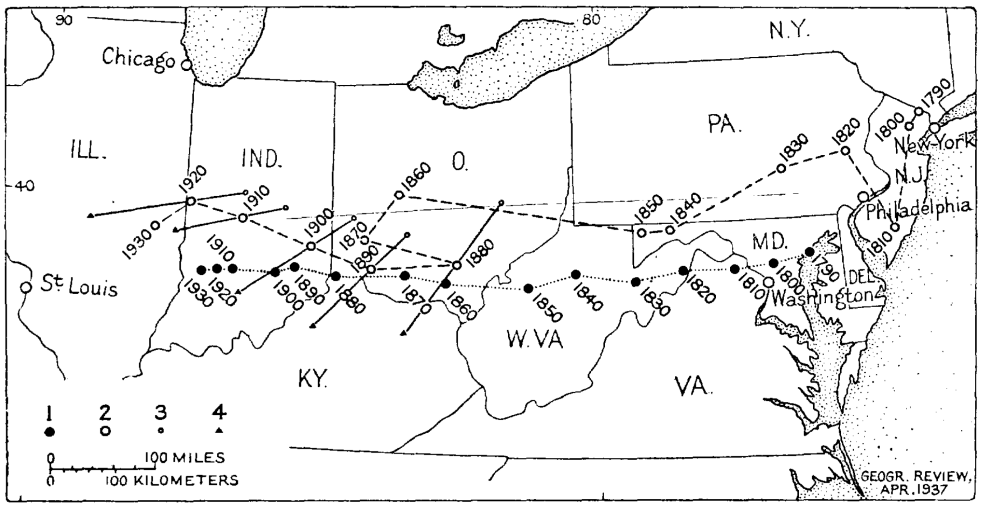


Fig. 6.3. Centrogramma del movimento del baricentro della popolazione e della formazione superiore negli Stati Uniti dal 1790 al 1930. Legenda: 1) popolazione generale, 2) educazione superiore, 3) popolazione maschile dell'educazione superiore, 4) popolazione femminile dell'educazione superiore. SVIATLOVSKY, EELLS 1937, p. 244.

della popolazione, permette di coglierne elementi fondamentali, quali la direzione e la velocità (rappresentata come distanza o spazio tra intervalli cronologici) di spostamento. Si tratta di informazioni a carattere sintetico che si prestano alla lettura e all'analisi geografica delle forme di occupazione sociale dello spazio.

6.3 La dispersione spaziale

La "dispersione spaziale" può essere intesa come uno degli attributi geografici più importanti di qualsiasi fenomeno traducibile a livello cartografico in una mappa di distribuzione per punti. A sua volta, la dispersione costituisce il principale indice di struttura di una maglia di stanziamento[20]. Per questo motivo si capisce come, dagli inizi degli anni '50 fino ai giorni nostri, si sia cercato di migliorare in modo sempre più rigoroso e critico le qualità formali del *point pattern analysis*. Non solo grazie alla definizione di nuovi metodi – come

[20] KARIEL 1970, p. 124: «Geographers interested in describing and analyzing settlement patterns have been concerned with trying to distinguish among different observed patterns. Although they have been able to describe them as dispersed, random or clusterd, until recently they have had no way of deciding objectively which of these types of patterns prevailed in a given area.»

la funzione K di Ripley – ma anche al miglioramento da parte della comunità scientifica di nuovi applicativi e programmi dedicati.

Per dispersione spaziale si intende il modo (in termini puramente topologici) secondo il quale gli oggetti o gli eventi di un determinato fenomeno sono disposti a livello spaziale: ordinato, casuale, concentrato (vedi figura 6.4). In altri termini, la disposizione, l'ordine, o la caratterizzazione distributiva che i punti – utilizzati per rappresentare un certo tipo di oggetto – assumono in termini spaziali. Come già indicato nei capitoli introduttivi di questo volume, la dispersione spaziale deve essere considerata come una delle caratteristiche materiali salienti di qualsiasi fenomeno reale: galassie, stelle, città, villaggi, case sparse, cellule o atomi. Ogni fenomeno reale possiede una propria logica spaziale. Solitamente la struttura spaziale non è altro che il risultato dell'interazione della realtà fisica, l'ambiente, la storia del fenomeno stesso (una sorta di *feedback*), con la propria logica o natura spaziale.

Occorre ricordare che l'interesse della geografia per le caratteristiche distributive va distinto in due aspetti principali. Da una parte si desidera indagare sulle caratteristiche distributive dell'insediamento umano perché da questo attributo è possibile desumere alcune caratteristiche funzionali della maglia di stanziamento. A seconda di come gli oggetti sono distribuiti in termini spaziali, essi tendono ad interagire in un modo piuttosto che in un altro. Dall'altra parte va ricordato che il geografo è interessato ad identificare indici di dispersione proprio perché questi consentono (come d'altronde qualsiasi forma di misurazione formale e rigorosa) di compiere confronti tra contesti regionali diversi[21].

A differenza di misure come la densità, che può essere considerata una misura di composizione, l'indice di dispersione, inteso come indicatore di struttura, costituisce un valore capace di distinguere in modo netto una rete di distribuzione dall'altra[22]. Se si prende in considerazione la figura 6.4, si possono immaginare le difficoltà che si potrebbero incontrare nel confrontare due segmenti adiacenti; o peggio ancora nel descrivere le caratteristiche distributive. È qui che tale tipo di studio o approccio quantitativo diviene utile allo studioso dei sistemi di stanziamento e dei paesaggi umani. La natura caratterizzante dell'indice di dispersione spaziale è superiore perché si riferisce appunto alla struttura della maglia di stanziamento; cioè come essa si vada a strutturare in termini spaziali e non tanto da quanti elementi essa sia composta. Ovverosia, il modo in cui gli elementi costituenti siano organizzati a livello geografico.

[21] Vedi paragrafo 1.6.

[22] Per una breve panoramica sul tema delle strutture distributive dell'insediamento umano consultare Hammond, McCullagh 1975, pp. 237-238.

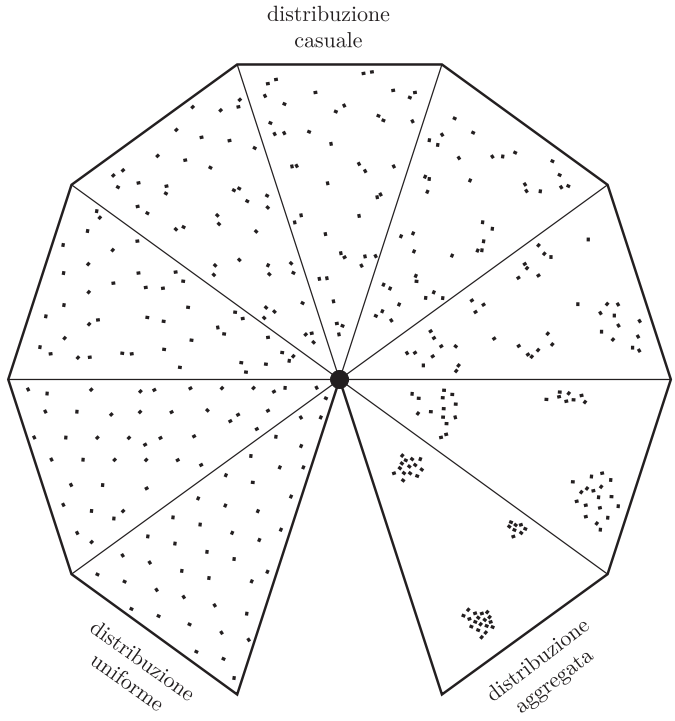


Fig. 6.4. Come si può osservare in questo grafico, i sistemi di distribuzione per punti possono essere classificati secondo una scala che, partendo da una dispersione quasi perfetta, arriva a una distribuzione concentrata. Il punto intermedio di questa sequenza è solitamente rappresentato da una distribuzione arbitraria. Il grafico vuole segnalare come le differenze tra le due estremità possano diventare impercettibili all'occhio se ci si avvicina ai valori centrali.

La dispersione spaziale non è solo un attributo; è anche una qualità che può essere quantificata in modo preciso. Vale a dire che, benché apparentemente caotica, essa può essere misurata.

Ciò ha delle implicazioni molto importanti in ambiti disciplinari quali la geografia umana, la geografia storica, la storia degli insediamenti e l'archeologia dei paesaggi storici. Il fatto che indicatori di struttura possano essere misurati significa che due maglie di stanziamento possano essere comparate, per l'appunto, a livello strutturale. L'essenza intrinseca della maglia di distribuzione viene così decodificata e tradotta a un linguaggio formale. Il concetto di dispersione acquista un significato ben specifico.

La possibilità di misurare la dispersione significa dunque poter procedere anche ad un raffronto formale tra due maglie di stanziamento. Conoscere que-

sto indice per due reti puntuali significa poter conoscere, ad esempio, quale delle due abbia la struttura più dispersa. Questo ragionamento acquista ancora maggiore rilevanza alla luce della storia e della teoria dei modelli rurali di stanziamento[23].

Riassumendo quello fin qui esposto, si può affermare che l'indice di dispersione spaziale rappresenta una scala numerica di struttura. Scala che indica valori che si collocano tra livelli di concentrazione e dispersione. A metà della scala si pone solitamente una distribuzione di tipo casuale. La figura 6.4 mostra le diverse strutture che una distribuzione per punti può assumere.

L'indice di dispersione acquista il suo nome dal livello massimo di organizzazione all'interno di questa scala. Il concetto di dispersione può essere associato anche a quello di ordine. Di solito, una rete ordinata presenta livelli di dispersione molto elevati. In effetti, l'indice massimo di dispersione può essere raggiunto all'interno di una rete puntuale perfettamente ordinata[24].

Un fattore molto importante da considerare è che le maglie di distribuzione per punti arbitrarie[25] sono un evento assai raro e inconsueto all'interno della realtà geografica. Per verificare l'effettiva natura casuale di una rete puntuale, è necessario che questa sia prodotta o generata al di fuori di qualsiasi forma di condizionamento o influenza. Cosa naturalmente molto difficile quando si parla della rete di un insediamento umano e della sua strutturazione.

È ovvio che i sistemi culturali, sociali, politici ed economici – quelli alla base delle forme di stanziamento umano – non possono che esercitare forti elementi di condizionamento. Sono questi sistemi in definitiva quelli che generano dei *patterns* spaziali. Ragione per la quale diviene molto difficile immaginare o ipotizzare una rete di distribuzione umana come veramente arbitraria o casuale.

Ogni rete di stanziamento umano conosciuta rappresenta perciò un'entità fisica che presenta valori di struttura spaziale che si trovano tra il raggruppamento assoluto e la casualità, o tra la casualità e la dispersione assoluta. Le reti di stanziamento umano possono apparire casuali anche se in realtà non lo sono perché, per definizione, esse non sono casuali. In termini formali, la casualità corrisponde all'assenza di qualsiasi forma di condizionamento. Condizione impossibile all'interno della realtà fisica o geografica. Il problema è che la maggior parte delle reti di stanziamento presenta strutture che spes-

[23] Ad esempio Hudson 1969.

[24] Per quanto riguarda il concetto di "ordine" in campo geocartografico fare riferimento a Sibley 1981, pp.1-3.

[25] A propostio dell'uso del termine *random* in geografia si rimanda a Getis 1977, pp. 59-61: «Geographers resort the term "random" when complex processes give rise to patterns which seem to have no systematic elements within them; that is, when the patterns exhibit no obvious cluster or obvious dispersal.»

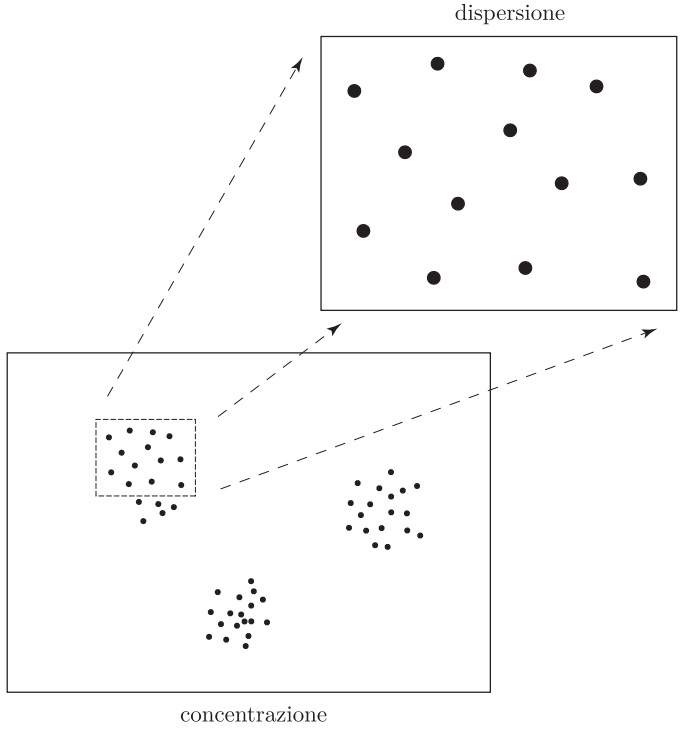


Fig. 6.5. I sistemi di distribuzione spaziale possiedono un livello di dispersione. Data una certa maglia di stanziamento, va tenuto presente che questo valore non è assoluto ma può variare con la scala geografica. La figura illustra un'area generale che presenta una rete di punti caratterizzata da un livello di concentrazione marcato. Se però si riduce l'area d'indagine al riguardo piccolo, si nota che quella porzione della maglia presenta una struttura dispersa.

so si avvicinano alla casualità (o così sembra), e spesso, proprio per il fatto che appaiono arbitrarie all'occhio umano, esse non vengono studiate a livello strutturale. L'analisi spaziale invece fornisce al ricercatore metodi specifici per lo studio approfondito a livello strutturale delle reti di distribuzione puntuale.

6.3.1 Dispersione spaziale e scala cartografica

La dispersione spaziale costituisce sicuramente uno degli elementi distintivi di qualsiasi sistema di stanziamento umano. Esso però non deve essere per nessun motivo concepito come un valore assoluto. Questo è stato uno dei principali errori della prima tradizione dei geografi che si sono cimentati nell'applicazione e nella definizione delle metodologie tradizionali.

Il problema potrebbe essere sintetizzato in questi termini: la dispersione spaziale di una maglia di stanziamento costituisce un suo tratto particolare e perciò identificativo; va tenuto sempre presente che tale attributo cambia (o può cambiare) a seconda della scala geografica di misurazione.

Lo studioso interessato a misurare la struttura spaziale di un sistema di stanziamento o di qualsiasi fenomeno sociale deve tener conto che l'oggetto in questione, se considerato su scale cartografiche diverse, potrebbe presentare caratteristiche relative alla dispersione spaziale anche contrastanti (vedi figura 6.5). Come verrà illustrato di seguito, la storia del *point pattern analysis* è stato un processo nei cui primi anni si sono identificati gli approcci fondamentali di misurazione; è stato poi come effetto di un ruolo crescente della scala cartografica in senso generale, che si sono adottate quelle misure che tengono conto anche delle differenze distributive in relazione alla variazione della scala.

6.4 Analisi a quadrati

Uno dei metodi più elementari, capaci di identificare le caratteristiche strutturali di una rete di distribuzione puntuale, è «l'analisi a quadrati». In inglese, questo metodo viene solitamente definito *quadrat analysis* e risale agli anni '50, periodo in cui ricercatori appartenenti alla botanica cercavano di formulare principi formali attraverso i quali giungere alla modellazione di processi stocastici[26].

A livello applicativo, l'analisi a quadrati non richiede particolare impegno e nello stesso modo non rappresenta un enorme carico a livello computazionale. È stata proprio questa facilità in termini applicativi ad avere permesso una larga diffusione in altre discipline[27]. In larga misura l'analisi a quadrati si fonda sull'idea dell'identificazione della variabilità delle frequenze. Infatti, questo sistema appartiene alla categoria dell'approccio a frequenze[28].

L'analisi a quadrati si basa sulla suddivisione dell'area di studio in tante tessere, generalmente quadrati di uguale dimensione[29]. Non è necessario che queste unità siano effettivamente di forma quadrata.

[26] David, Moore 1954.

[27] Sull'applicazione in geografia consultare Unwin 1986, pp. 92.

[28] Consultare Harvey 1966, pp. 82-83. Non a caso il calcolo dell'indice di dispersione realizzato con questo metodo si basa proprio sulla varianza campionaria.

[29] O'Sullivan, Unwin 2003, p. 84.

A dispetto del nome di questa analisi, si può frazionare il territorio interessato con tessere anche rettangolari o triangolari. Le uniche condizioni sono che le unità siano tutte della stessa estensione e che la griglia che queste compongono vada a coprire l'intera superficie di studio. Per ovvii motivi, il metodo più semplice è quello della suddivisione in tessere quadrate[30].

Una volta coperta la rete di distribuzione puntuale con la griglia di frazionamento, si procede semplicemente con la classificazione o la costruzione dell'istogramma di frequenza. Benché l'istogramma non costituisca l'analisi in sé, l'esempio illustrato nella figura 6.6 dimostra come le differenze dei *patterns* si riflettano immediatamente sui valori di frequenza.

È facile intuire come i dati raccolti con procedure come quelle illustrate dalla figura 6.6 possano essere già sufficienti per essere sfruttate empiricamente in modo da ottenere indicazione di struttura spaziale per mappe di distribuzione per punti. Il primo caso (A) presenta una struttura con un elevato grado di dispersione: la distribuzione dei punti, in effetti, appare molto regolare. L'istogramma a fianco dell'esempio citato presenta le classi organizzate secondo la frequenza nelle diverse celle. Si può osservare perciò che la caratteristica che contraddistingue un elevato grado di dispersione è il fatto che i dati tendono a convergere in poche classi. In questo caso le classi con 0 o prevalentemente 1 punto al loro interno.

Al contrario del primo caso, l'esempio B presenta una rete di distribuzione accentrata. L'istogramma a fianco presenta i dati osservati. Si può osservare come, nel caso di reti aggregate, l'istogramma prodotto si caratterizzi per il contrasto di molte celle con 0 punti e poche unità con un livello elevato di punti. All'analisi dei quadrati la rete di distribuzione tende perciò a presentare il raggruppamento dei punti in una quantità limitata di celle. Di riflesso, il resto delle celle apparirà vuoto.

Questo esempio ci permette di capire come anche indicazioni di frequenza possano essere chiarificatrici in termini qualitativi. L'analisi a quadrati offre però elementi fondamentali anche in termini statistici. Dunque misurazioni che possono essere sfruttate per l'identificazione di valori o indici precisi.

La scala di dispersione calcolata con il metodo dei quadrati va intesa come una misura che permette di cogliere alcuni elementi della struttura spaziale sulla base dei valori di frequenza. L'indice di dispersione in questione è una scala numerica organizzata con valori che, partendo da 0, passano da 1 e arrivano a quote superiori a 1 e, in linea del tutto teorica, all'infinito. Anche se non esiste un modo ortodosso per definire tale valore, per motivi di praticità

[30] Una trattazione dettagliata della *quadrat analysis* si trova in Silk 1979, pp. 95-105.

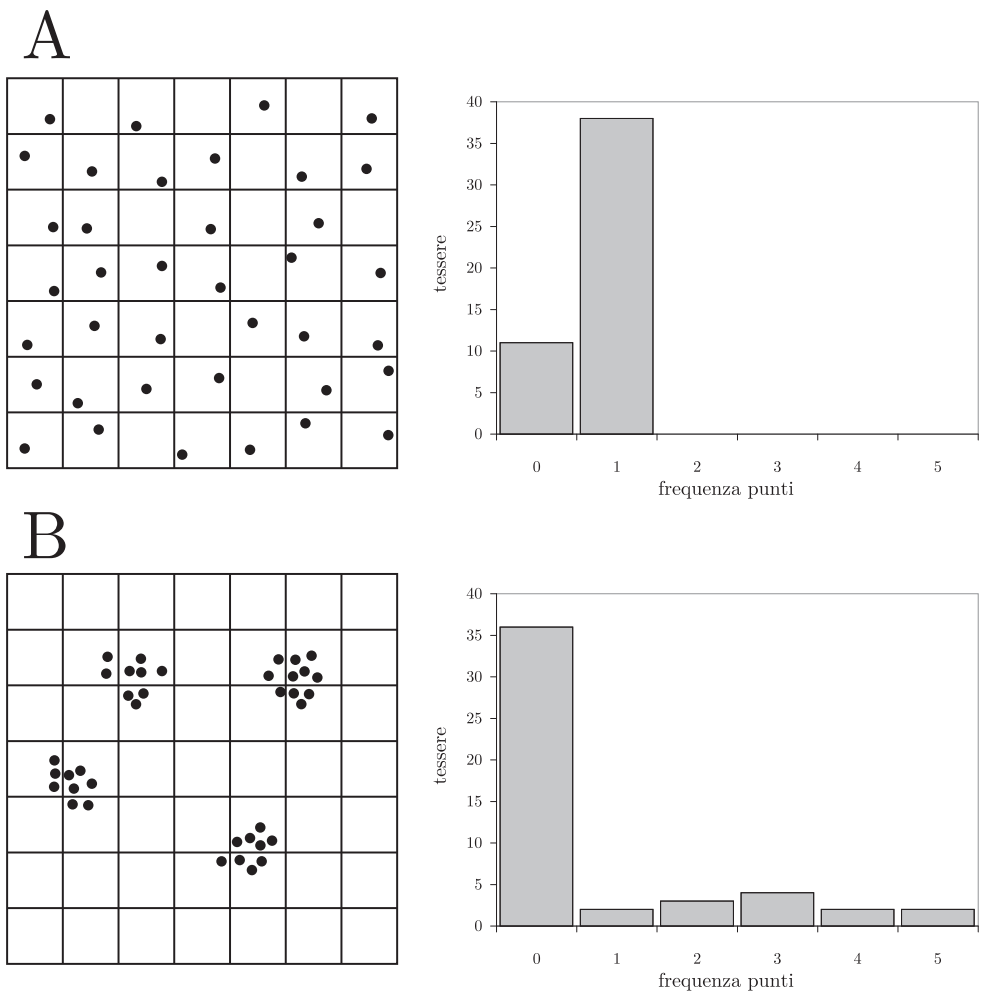


Fig. 6.6. Confronto tra due esempi di istogrammi di distribuzione relativi a due carte per punti. Si può osservare come le differenze della struttura spaziale presentino a livello statistico schemi riassuntivi molto diversi.

chiameremo l'indice di dispersione calcolato con il metodo dei quadrati come ICS[31].

[31] La notazione ICS, o *index of clumping*, fu introdotta da Douglas 1975

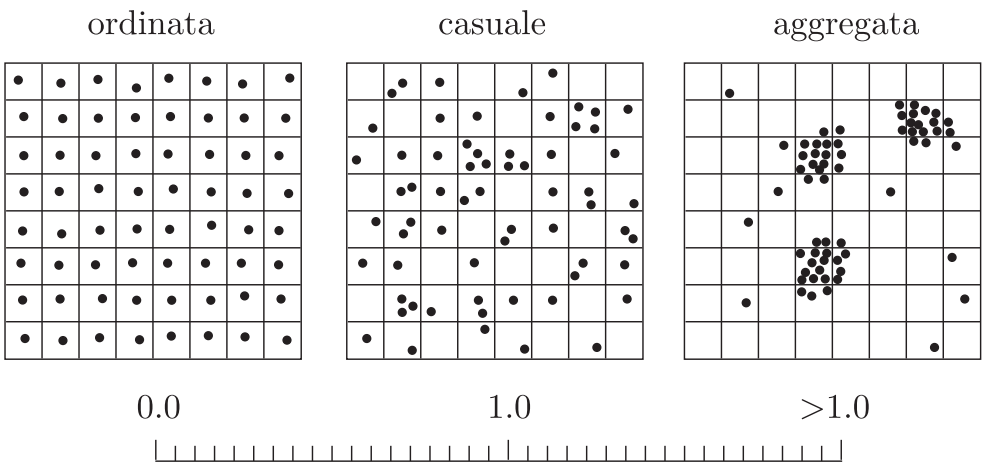


Fig. 6.7. La scala di dispersione (ICS) per la valutazione dei risultati ottenuti con il metodo dei quadrati. Come si può osservare dalla rappresentazione della scala, il metodo dei quadrati rileva il valore 1 per i sistemi di distribuzione casuale, valori inferiori a 1 per reti disperse o ordinate; infine valori superiori a 1 per reti aggregate.

Prima di illustrare la metodologia di calcolo diremo che, quando ICS sia vicino a 0 o, in ogni caso, inferiore a 1, si parla di una rete per punti dispersa; anche se, nel caso specifico di questo tipo d'analisi, sarebbe meglio definirla "ordinata". Se invece ICS presenta un valore uguale o vicino a 1, si può concludere che si tratta di una rete distributiva casuale. Infine, se il valore rilevato è superiore a 1, si tratta di una rete aggregata.

La formula per il calcolo dell'indice di dispersione con le analisi a quadrati è:

$$ICS = \frac{s^2}{\bar{x} - 1} \tag{6.4}$$

dove

$$s^2 = \frac{\sum_{j=1}^{n} (x_i - \bar{x})^2}{(n-1)} \tag{6.5}$$

ovverosia, l'ICS può essere inteso come un generico rapporto tra la varianza e la media delle frequenze registrate nelle tessere. Il procedimento può essere

realizzato calcolando media e varianza delle tessere nel loro complesso con le formule appena descritte. L'indice di dispersione calcolato secondo il metodo dei quadrati può essere interpretato perciò come:

$$ICS = \frac{\text{varianza punti nelle tessere}}{\text{media punti nelle tessere}}$$

Si prendano in considerazione gli esempi illustrati nella figura 6.6. L'esempio trattato si riferisce al calcolo della dispersione per il caso A. I dati osservati sono dunque:

frequenza punti x	tessere p
0	11
1	38
2	0
3	0
4	0
5	0
6	0

Tabella 6.3. Frequenza dei punti per celle dell'esempio A. Vedi figura 6.6

Per prima cosa, si procede dunque al calcolo della media dei punti per cella con la formula della media ponderata:

$$\bar{x}_A = \frac{\sum_{j=1}^{n} x_i \times p_i}{\sum_{j=1}^{n} p_i}$$

cioè

$$\bar{x}_A = \frac{(0 \times 11) + (1 \times 38)}{11 + 38} = 0.775$$

A sua volta, la varianza per i dati illustrati va calcolata secondo 6.5:

$$s^2{}_A = \frac{((0 - \bar{x}_A)^2 \times 11) + ((1 - \bar{x}_A)^2 \times 38)}{49 - 1} = 0.1777$$

L'indice di dispersione calcolato con l'analisi a quadrati per il caso A corrisponde dunque a:

$$ICS_A = \frac{s^2}{\bar{x}} = \frac{0.1777}{0.775} = 0.229$$

che risulta essere un valore inferiore a 1 particolarmente basso e dunque riferibile a una rete di punti caratterizzata da un distribuzione dispersa particolarmente marcata.

La seguente tabella presenta invece i dati relativi al caso B della figura 6.6.

frequenza punti x	tessere p
0	36
1	2
2	3
3	4
4	2
5	2
6	0

Tabella 6.4. Frequenza dei punti per celle nell'esempio B. Vedi figura 6.6

$$\bar{x}_B = \frac{(0 \times 36) + (1 \times 2) + \cdots + (5 \times 2)}{36 + 2 + 3 + 4 + 2 + 2} = 0.775$$

mentre

$${s^2}_B = \frac{((1 - \bar{x}_B)^2 \times 36) + ((1 - \bar{x}_B)^2 \times 2) \cdots (5 - \bar{x}_B)^2 \times 2)}{49 - 1} = 2.136$$

L'indice di dispersione viene così stabilito in:

$$ICS_B = \frac{s^2}{\bar{x}} = \frac{2.136}{0.775} = 2.754$$

Nel caso B perciò viene identificato un valore al di sopra di 1, già segnalato come indicatore di una dispersione aggregata.

6.4.1 Limiti dell'analisi a quadrati

Nella storia di questo metodo diverse critiche sono state mosse. Se da una parte hanno permesso di migliorare ed espandere la sua applicabilità in termini di modellazione dei dati, nello stesso modo queste critiche – che evidenziano

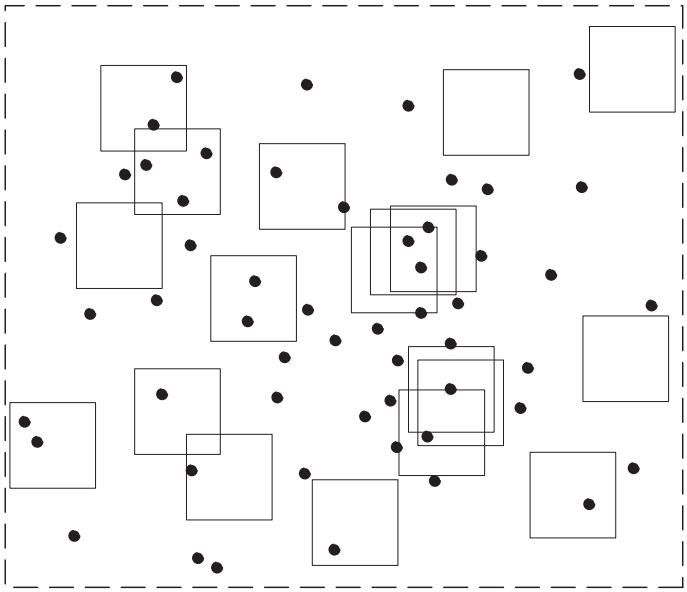


Fig. 6.8. Molti limiti del metodo di misurazione a quadrati possono essere risolti con la misurazione continua delle frequenze dei punti grazie ad una tessera posizionata casualmente.

effettivi limiti applicativi – hanno determinato il progressivo sviluppo e la diffusione di altri metodi.

L'applicazione dell'analisi a quadrati presenta alcuni inconvenienti. In particolare a questo tipo d'analisi sono state mosse due obiezioni. La prima si riferisce alla dimensione ideale delle tessere. È chiaro che si tratta di un fattore che può trasformare in modo più o meno rilevante i risultati delle analisi.

Al problema delle dimensioni viene solitamente applicata una soluzione abbastanza empirica. Da tempo è stato dimostrato come per questo tipo d'analisi il maggior grado di significatività dei risultati nel processo di classificazione spaziale per mezzo di una griglia possa essere raggiunto quando, per ogni tessera dell'area analizzata, ci sono in media due punti. Per questo motivo, per individuare l'area ideale di una tessera basta applicare la seguente formula:

$$a = 2 \times \frac{A}{n} \tag{6.6}$$

dove a è l'area ideale della tessera, A quella della superficie di studio e n il numero di punti nella maglia. Se invece si vuole conoscere la dimensione ideale del lato l di una tessera quadrata, basta calcolare la radice quadrata dell'area ideale:

$$l = \sqrt{2 \times \frac{A}{n}} \tag{6.7}$$

La seconda obiezione mossa a questo tipo d'analisi è quella dell'origine della griglia. In effetti, di fronte a una stessa rete di distribuzione, la stessa griglia produrrà risultati diversi a seconda della sua posizione.

Una soluzione rigorosa al problema dell'origine è quella che prevede innanzitutto l'individuazione dell'area ideale della cella secondo la formula 6.6. Successivamente si procede a far cadere casualmente (come illustrato nella figura 6.8) sull'area di studio la tessera e registrare su una tabella la classe di frequenza (il numero di punti all'interno della cella). Successivamente si procede a ripetere questa operazione più volte. Si possono in questo modo raccogliere più dati di quelli che avrebbe prodotto una griglia tradizionale. In effetti, con questo metodo, i punti potrebbero essere contati più volte ma l'incremento dei dati produrrebbe in ogni caso dati statistici più attendibili[32]. Naturalmente nel processo di calcolo statistico dell'indice ICS si dovrà tenere conto del numero di volte che l'operazione di misurazione è stata realizzata.

6.5 Il vicino più prossimo

Gli inconvenienti legati ai metodi di frequenza determinarono il progressivo sviluppo e la diffusione dei metodi basati sulle "distanze". Tra questi, quello ad avere goduto di maggiore popolarità nell'ambito della geografia umana è stato quello normalmente definito come "vicino più prossimo". Metodo introdotto nella geografia da Dacey sul finire degli anni '50, benché anche questo metodo traesse le sue origini sempre dall'ecologia[33].

Si tratta di un metodo per il calcolo dell'indice di dispersione in una rete di distribuzione per punti. Il metodo è abbastanza semplice: consiste nel confronto dei valori delle distanze osservate con quelle attese, cioè quelle relative al modello di distribuzione casuale di Poisson. Come verrà spiegato di seguito, il calcolo del vicino prossimo, benché concettualmente molto semplice, richiede (soprattutto nel caso di analisi riferite a reti molto grandi di punti) l'ausilio di un calcolatore elettronico; sia per motivi di carico di lavoro, sia perché è necessario identificare con precisione quale sia effettivamente il vicino più prossimo.

[32] O'Sullivan, Unwin 2003, p. 82.

[33] Il metodo è stato introdotto in ambito geografico da Dacey 1958 e Dacey 1960. Il metodo era stato sviluppato in origine da Clark, Evans 1954.

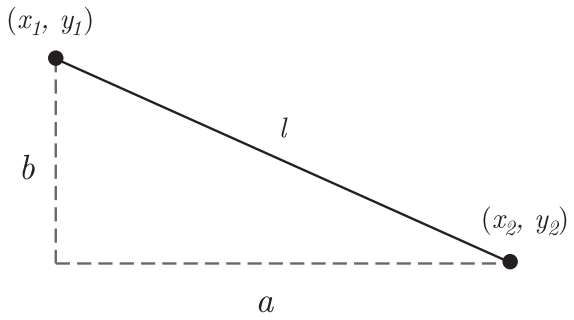


Fig. 6.9. Il metodo del vicino prossimo si basa a livello pratico sul calcolo delle distanze tra insediamenti. La distanza di due punti su un piano può essere calcolata facilmente grazie al teorema di Pitagora.

Il primo passo per la misurazione dell'indice di dispersione con il metodo del vicino prossimo richiede l'identificazione del vicino più prossimo[34] per ogni elemento di una rete di distribuzione a punti.

Ogni segmento di distanza tra ogni punto ed il suo vicino più prossimo viene rilevato e registrato. Non è difficile immaginare quanto potesse essere lungo e dispendioso compiere tali operazioni in assenza di un calcolatore. Le applicazioni di calcolo per la misurazione della distanza fra due punti si basano sul teorema di Pitagora ed utilizzano la formula canonica[35]:

$$l = \sqrt[2]{a^2 + b^2} \tag{6.8}$$

dove a e b rappresentano due lati noti di un triangolo e l la lunghezza dell'ipotenusa. Dati dunque due punti rappresentati dalle coordinate (x_1, y_1) e (x_2, y_2), le dimensioni di a e b possono essere immaginate come la differenze[36] rispettivamente tra le ascisse e le ordinate:

$$a = x_1 - x_2$$

e

$$b = y_1 - y_2$$

[34] Unwin 1986, pp. 98-99.

[35] O'Sulliva, Unwin 2003, pp. 34-35; inoltre consultare Fotheringham, Brunsdon, Charlton 2000. pp. 20-21.

[36] Dato che nel calcolo dell'ipotenusa tali differenze vengono elevate al quadrato, non interessa se le differenze risultano negative o positive; in altre parole $(10 - 6)^2 = (6 - 10)^2$.

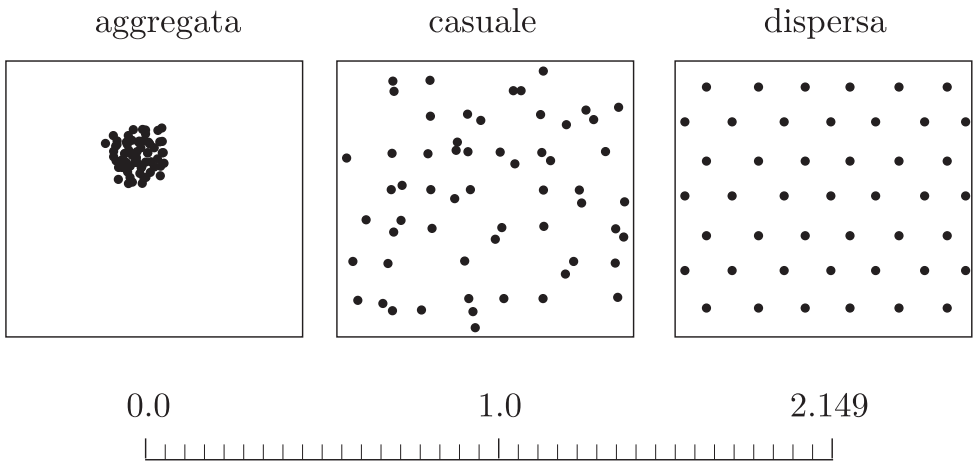


Fig. 6.10. Scala dell'indice R di distribuzione basata sull'analisi del vicino prossimo. Fare attenzione al fatto che la direzione dell'indice R è inversa a quella dell'indice ICS. Infatti lo 0 equivale ad un sistema per punti dove la concentrazione è tale che le distanze fra i punti sono azzerate. Un'altra differenza è che l'indice R presenta un valore massimo assoluto, 2.149, che si registra quando la maglia dei punti è distribuita secondo la rete di massima efficienza.

Dunque, estendendo la formula 6.8, ai fini del calcolo della distanza tra due punti espressi in termini di coordinate cartesiane, si ottiene:

$$l = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{6.9}$$

Fare attenzione al fatto che la distanza dal vicino più prossimo (l_i) viene calcolata per ogni singolo punto. Vale a dire che, se due punti sono reciprocamente vicini, si dovrà tenere conto due volte della distanza che li separa. Così ad esempio, data una rete di distribuzione di 58 punti, dovranno essere misurate 58 distanze ai rispettivi vicini prossimi. Dal calcolo di tutte le distanze si otterrà poi una distanza media osservata; questa viene solitamente definita r_o. Come è stato già suggerito, r_o corrisponde dunque alla media di l_i. Perciò:

$$r_o = \bar{l} = \frac{\sum_{i=1}^{n} l_i}{n} \tag{6.10}$$

o, più semplicemente, la media di tutti i segmenti o distanze da ogni singolo punto al vicino più prossimo.

	coordinate			*coordinate*		
punto	x	y	vicino	x	y	l
1	1691983	4815218	2	1691866	4812528	2693
2	1691866	4812528	1	1691983	4815218	2693
3	1696564	4815383	1	1691983	4815218	4584
4	1686039	4810052	31	1683819	4807490	3390
5	1696671	4763285	61	1699958	4765084	3747
6	1739241	4751990	7	1736607	4755214	4163
7	1736607	4755214	6	1739241	4751990	4163
8	1735086	4749563	9	1730530	4749461	4557
9	1730530	4749461	8	1735086	4749563	4557
10	1705821	4762984	61	1699958	4765084	6228
11	1725369	4769393	27	1726334	4768684	1197
12	1720553	4770129	13	1721116	4774250	4159
13	1721116	4774250	59	1719250	4776329	2794
14	1725087	4772440	53	1726334	4774848	2712
15	1733183	4774012	18	1733163	4777519	3507
16	1729003	4775703	53	1726334	4774848	2803
17	1727574	4782855	55	1725114	4783623	2577
18	1733163	4777519	15	1733183	4774012	3507
19	1735115	4770703	15	1733183	4774012	3832
20	1734588	4765009	22	1736870	4760631	4937
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
80	1674296	4804415	46	1676288	4799943	4896
81	1672886	4797461	46	1676288	4799943	4211
82	1681171	4795479	48	1681729	4794804	876
83	1677989	4797365	46	1676288	4799943	3089
84	1677365	4794314	83	1677989	4797365	3114
85	1676278	4786735	47	1676780	4784920	1883
85	1676278	4786735	47	1676780	4784920	1883
86	1666692	4796282	88	1666365	4791752	4542
87	1665627	4800761	86	1666692	4796282	4604
88	1666365	4791752	86	1666692	4796282	4542
89	1684048	4778132	43	1682900	4782846	4852
90	1672542	4775177	47	1676780	4784920	10625
91	1668683	4785523	88	1666365	4791752	6646
92	1673687	4814875	76	1671687	4809950	5316

Tabella 6.5. Tabella relativa alle coordinate Gauss-Boaga delle pievi citate nelle *Seconde Decime* (Giusti, Guidi 1942) con l'identificazione del relativo vicino ed il calcolo della distanza (l).

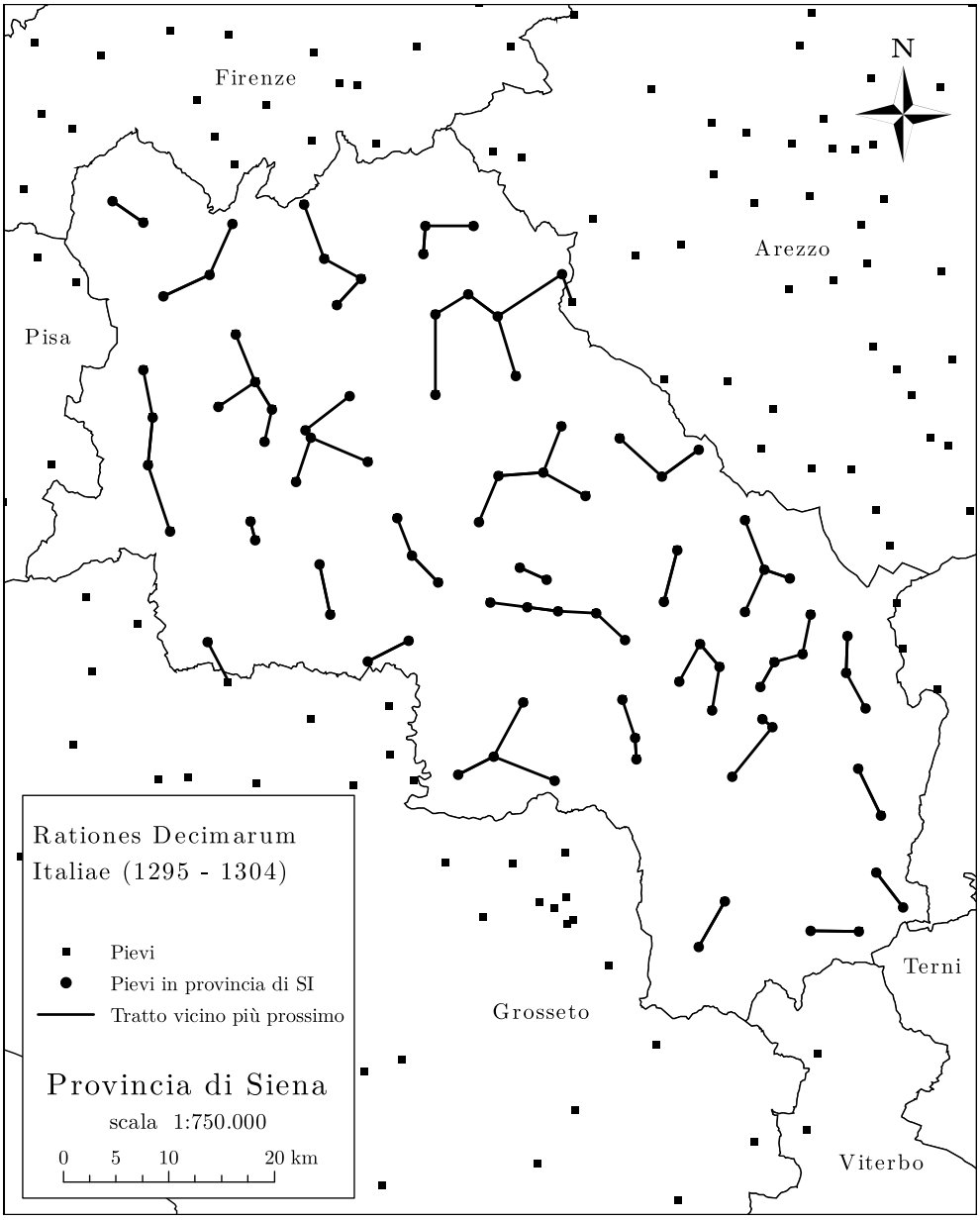


Fig. 6.11. Rappresentazione cartografica di tutti i segmenti che collegano ogni punto al suo vicino prossimo per le pievi del territorio della provincia di Siena citate nelle *Seconde Decime*. Il metodo del vicino più prossimo si basa sul calcolo della distanza che separa ogni punto da quello immediatamente più vicino. L'operazione d'identificazione va realizzata per ogni punto, anche nel caso la vicinanza prossima sia reciproca.

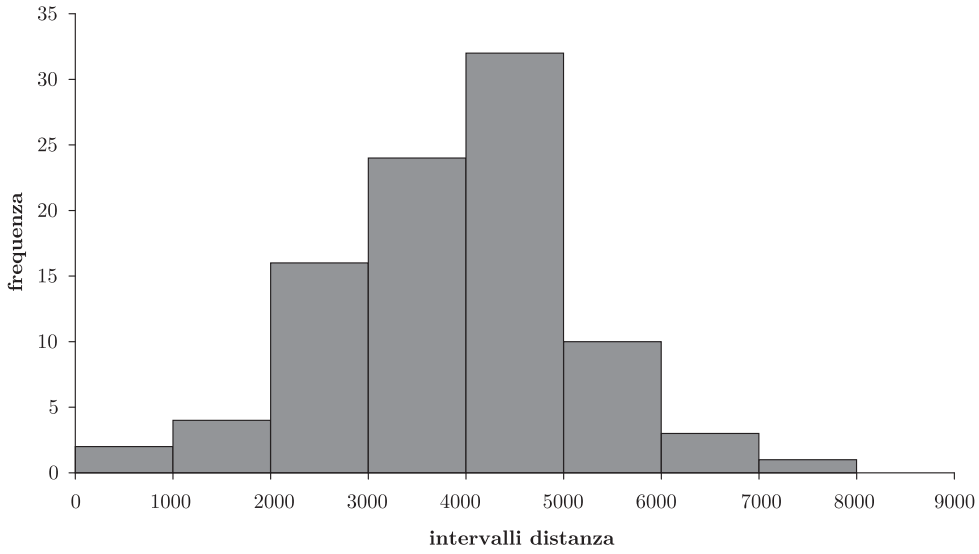


Fig. 6.12. Istogrammi delle distanze ai vicini più prossimi per le pievi della provincia di Siena riportate nella tabella 6.5.

La seconda parte relativa al calcolo dell'indice di dispersione R va riferita al computo della distanza media attesa (r_e) per una rete di punti casuale.

Il calcolo di r_e risponde alla media attesa al vicino più prossimo per una rete di punti casuale o di Poisson. In pratica, conoscendo un'area determinata A e definendo un numero noto di punti n, è possibile modellare (e pronosticare dunque) la media della distanza tra i punti e i propri vicini più prossimi. Questa media va indicata con il simbolo r_e e viene solitamente calcolata con[37]:

$$r_e = \frac{1}{2 \times \sqrt{\lambda}} \tag{6.11}$$

dove $\lambda = n/A$, ovvero la densità o frequenza di punti per una distribuzione di Poisson.

Questo aspetto è molto importante dato che, per potere calcolare effettivamente l'indice di dispersione R, a differenza dell'analisi a quadrati, è fondamentale conoscere l'estensione (in termini quantitativi) del territorio o campione indagato (A). Vale a dire che l'area d'indagine fa parte dei requisiti di questo tipo d'analisi. Fattore che, come sarà esposto di seguito, comporta non pochi problemi.

[37] Clark, Evans 1954, p. 447.

A partire da r_o e r_e è poi possibile giungere all'identificazione dell'indice di dispersione R con:

$$R = \frac{r_o}{r_e} \tag{6.12}$$

L'indice di dispersione R del vicino prossimo permette di calcolare con precisione le tendenze e il grado di dispersione di un sistema di stanziamento. L'indice R va inteso come una scala che presenta valori vicini allo 0 per reti puntuali aggregate, vicini a 1 per quelle casuali e a 2.149 come valore limite per una rete dispersa (vedi figura 6.10). Questo valore limite di dispersione è riferito ad una rete puntuale organizzata su una griglia perfetta di triangoli equilateri. Difatti si può considerare la disposizione su di una griglia triangolare il modo più coerente ed efficiente per organizzare delle entità nello spazio[38].

Ad esempio, la distanza media osservata del vicino più prossimo (r_o) per le pievi medievali nella provincia di Siena è di 3.94 chilometri. All'interno della struttura spaziale della maglia, le pievi sono poste cioè mediamente a questa distanza l'una dall'altra. I dati raccolti possono essere raggruppati e sintetizzati in un istogramma (vedi figura 6.12). Questo può dare delle indicazioni generali sul grado di dispersione della rete di distribuzione qualora si abbia una chiara percezione della scala geografica. Di fatto, l'istogramma delle r_o può anche essere sfruttato come strumento di descrizione del fenomeno in termini spaziali.

Per il calcolo della media attesa del vicino più prossimo delle pievi medievale nella Provincia di Siena è sufficiente procedere all'applicazione di 6.11 conoscendo i valori di n (92 pievi) e A (3820.64 km^2) per il calcolo di λ. Si ottiene così:

$$r_e = \frac{1}{2 \times \sqrt{92/3820.64}} = 3.21 km^2$$

Si registra un valore atteso inferiore a quello osservato. Questo significa che tendenzialmente le distanze osservate risultano essere superiori a quelle attese. Dunque l'indice di dispersione R delle pievi medievali nel territorio di Siena risulta essere:

$$R = \frac{3.94}{3.21} = 1.22$$

che, in base a quanto indicato sopra, risulta essere un valore associato con una distribuzione dispersa.

Nell'ambito dello studio dei paesaggi antropici è importante sottolineare, qualora ce ne fosse bisogno, l'importanza di questo tipo di misurazione in

[38] Vedi modello di Christaller.

termini geografici. Essa consente infatti di giungere a forme di comparazioni precise e puntuali. In altri termini, è possibile confrontare, ad esempio, l'indice di dispersione delle pievi del territorio senese con altre zone interessate dallo stesso fenomeno e, a partire dai risultati ottenuti, rilevare ed interpretare i motivi di tali differenze.

6.5.1 Test di significatività di *R*

Un altro aspetto fondamentale per il calcolo e l'applicazione dell'indice di dispersione R è il suo test statistico di significatività. Si tratta di una procedura di quantificazione che permette di stabilire con precisione e un certo grado di confidenza se i valori rilevati per R possono essere considerati casuali o meno. Sappiamo ad esempio che valori intorno a 1 corrispondono a una distribuzione casuale. Ma di quanto si deve discostare l'indice R da 1 per poter iniziare ad essere considerato espressione di una distribuzione dispersa (per valori superiori a 1) o aggregata (valori inferiori a 1)? Ad esempio, 0.91 deve essere considerato come indice di una rete aggregata o casuale? Cioè qual è la soglia esatta con la quale si passa dalla casualità a valori indicatori di strutture aggregate o disperse.

Il problema non ha una soluzione meccanica. In altri termini, non esiste una soglia predefinita, sopra 1 per valori dispersi e sotto 1 per valori aggregati. La soglia cambia e si avvicina sempre di più a 1 man mano che aumenta il numero di punti sotto esame. Oppure aumenta la loro frequenza per unità spaziale.

Fortunatamente la statistica offre un metodo molto semplice ed efficace denominato l'"errore standard" che può essere calcolato nel caso dell'indice del vicino più prossimo:

$$z = \frac{r_o - r_e}{\sigma r_o} \tag{6.13}$$

dove σr_o è definito come errore standard della frequenza dei valori studiati e viene calcolato con:

$$\sigma r_o = \frac{0.26136}{\sqrt{n\lambda}} \tag{6.14}$$

In sostanza la z fornisce un valore che indica un indice di significatività per il risultato dei dati osservati. Fornisce cioè indicazioni di quanto i valori osservati possano essere considerati ancora (o meno) di una rete di distribuzione casuale. Se il valore di z supera anche di poco le soglie prestabilite relative ai valori osservati, allora si può concludere che i valori non possono essere con-

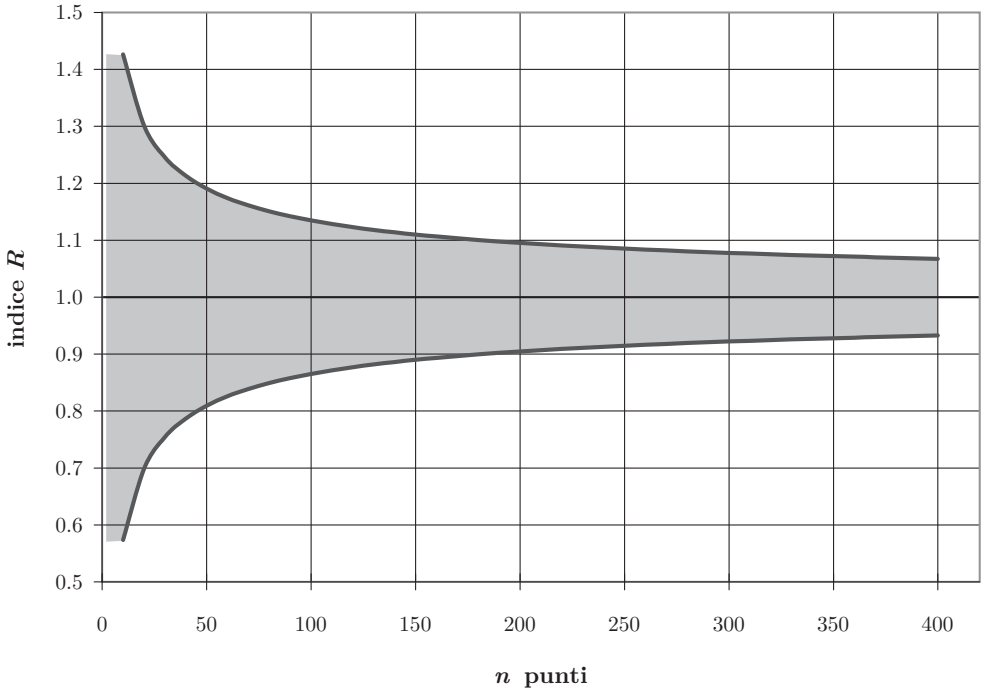


Fig. 6.13. Il calcolo dell'indice di dispersione R comporta una variazione della significatività con l'incremento dei punti analizzati. Questa variazione consiste in una progressiva diminuzione tra quelli che identificano una maglia casuale e una aggregata, o casuale e dispersa. In altra parole, l'indice R per una rete regolare apparirà sempre più vicino a 1 con l'incremento del numero dei punti analizzati. Nelle ascisse sono rappresentati i numeri di punti. Sulle ordinate invece è rappresentata la scala dell'indice R. L'area grigia del grafico presenta la zona che può essere considerata statisticamente come casuale.

siderati come relativi ad una rete casuale. Solitamente sono la soglia di 1.96 per un 95% di confidenza e 2.58 per un 99% di confidenza.

Si prenda in considerazione il caso delle pievi medievali della provincia di Siena. Applicando 6.13 si ottiene che:

$$z = \frac{3.94 - 3.21}{\frac{0.26136}{\sqrt{92 \times 0.024}}} = 4.14$$

In questo caso z supera la soglia di 2.58 che ci permette di confutare al 99% che la distribuzione illustrata nella figura 6.11 possa essere considerata una distribuzione di tipo casuale. Il test statistico permette di confermare che con un'elevata soglia di confidenza la distribuzione in questione è dispersa.

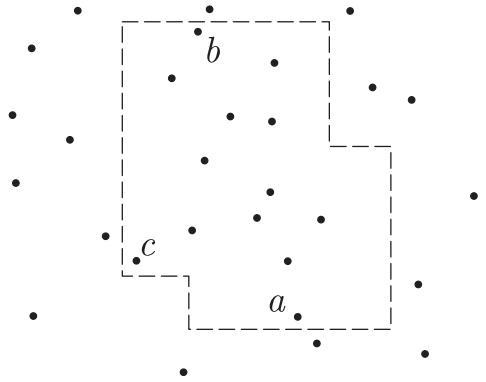


Fig. 6.14. Nell'applicazione del metodo del vicino prossimo bisogna considerare anche come i sistemi di stanziamento non siano delle entità perfettamente racchiuse all'interno delle aree di studio. Nel computo dei valori occorre tener presente che quasi sempre le maglie di stanziamento continuano al di là dei confini delle zone indagate.

La tabella 6.6 mostra il calcolo delle soglie limite per R e r_o sia accentrate e disperse per distribuzioni di punti da 10 a 400 unità. I valori di questa tabella sono stati calcolati a partire sempre dalla formula 6.13.

6.5.2 Limiti del vicino più prossimo

Nel tempo diversi sono stati i limiti che molti studiosi hanno rilevato in relazione all'applicazione del metodo del vicino più prossimo nell'ambito dello studio delle maglie di stanziamento.

Il limite più importante è quello dei confini dell'area di studio. Come è stato segnalato precedentemente, l'applicazione del metodo richiede il computo della superficie dell'area di studio. Il problema sorge quando alcuni punti all'interno dell'area di studio hanno i propri vicini al di fuori dei confini dell'area di studio. In questi casi, se il punto prende in considerazione esclusivamente vicini all'interno dall'area di studio si produce un'alterazione dell'indice R. Il problema non si può risolvere semplicemente allargando l'area di studio dato che questo comporterebbe nuovi casi come il precedente. Occorre perciò fare attenzione a questo aspetto e verificare, soprattutto nel processo di raccolta dati, di registrare anche siti esterni ai confini, ma tra i quali si pensa possano essere vicini di quelli studiati. Se i dati vengono raccolti correttamente, a livello computazionale non ci sono grandi problemi dato che gli algoritmi sanno tenere conto di questo limite del metodo.

				dispersa		*concentrata*	
n punti	λ	r_e	σ_{r_e}	$r_o > 1$	R	$r_o < 1$	R
10	0.1	1.5811	0.2614	2.2554	1.4265	0.9068	0.5735
20	0.2	1.1180	0.1307	1.4552	1.3016	0.7809	0.6984
30	0.3	0.9129	0.0871	1.1376	1.2462	0.6881	0.7538
40	0.4	0.7906	0.0653	0.9591	1.2132	0.6220	0.7868
50	0.5	0.7071	0.0523	0.8420	1.1907	0.5722	0.8093
60	0.6	0.6455	0.0436	0.7579	1.1741	0.5331	0.8259
70	0.7	0.5976	0.0373	0.6939	1.1612	0.5013	0.8388
80	0.8	0.5590	0.0327	0.6433	1.1508	0.4747	0.8492
90	0.9	0.5270	0.0290	0.6020	1.1422	0.4521	0.8578
100	1.0	0.5000	0.0261	0.5674	1.1349	0.4326	0.8651
110	1.1	0.4767	0.0238	0.5380	1.1286	0.4154	0.8714
120	1.2	0.4564	0.0218	0.5126	1.1231	0.4002	0.8769
130	1.3	0.4385	0.0201	0.4904	1.1183	0.3867	0.8817
140	1.4	0.4226	0.0187	0.4707	1.1140	0.3744	0.8860
150	1.5	0.4082	0.0174	0.4532	1.1101	0.3633	0.8899
160	1.6	0.3953	0.0163	0.4374	1.1066	0.3531	0.8934
170	1.7	0.3835	0.0154	0.4231	1.1034	0.3438	0.8966
180	1.8	0.3727	0.0145	0.4101	1.1005	0.3352	0.8995
190	1.9	0.3627	0.0138	0.3982	1.0978	0.3272	0.9022
200	2.0	0.3536	0.0131	0.3873	1.0954	0.3198	0.9046
210	2.1	0.3450	0.0124	0.3771	1.0931	0.3129	0.9069
220	2.2	0.3371	0.0119	0.3678	1.0909	0.3064	0.9091
230	2.3	0.3297	0.0114	0.3590	1.0889	0.3004	0.9111
240	2.4	0.3227	0.0109	0.3508	1.0871	0.2947	0.9129
250	2.5	0.3162	0.0105	0.3432	1.0853	0.2893	0.9147
260	2.6	0.3101	0.0101	0.3360	1.0836	0.2842	0.9164
270	2.7	0.3043	0.0097	0.3293	1.0821	0.2793	0.9179
280	2.8	0.2988	0.0093	0.3229	1.0806	0.2747	0.9194
290	2.9	0.2936	0.0090	0.3169	1.0792	0.2704	0.9208
300	3.0	0.2887	0.0087	0.3112	1.0779	0.2662	0.9221
320	3.2	0.2795	0.0082	0.3006	1.0754	0.2584	0.9246
340	3.4	0.2712	0.0077	0.2910	1.0731	0.2513	0.9269
350	3.5	0.2673	0.0075	0.2865	1.0721	0.2480	0.9279
360	3.6	0.2635	0.0073	0.2823	1.0711	0.2448	0.9289
380	3.8	0.2565	0.0069	0.2742	1.0692	0.2387	0.9308
400	4.0	0.2500	0.0065	0.2669	1.0674	0.2331	0.9326

Tabella 6.6. Intervallo di confidenza di R. Computo delle soglie minime e massime per il test statististici di conformità alla struttura casuale secondo la formula 6.13. Ad esempio con 90 punti se $R > 1.14$ allora non è possibile affermare cha la struttura sia casuale. In termini statistitici essa apparirà più conforme ad una struttura dispersa. Sempre con 90 punti $R < 0.8578$ rappresenta una soglia che indica valori aggregati.

6.6 La funzione K di Ripley

Come già più volte accennato, il *point pattern analysis* si è caratterizzato, fin dalla sua nascita, per un progressivo riconoscimento dei limiti che i metodi – via via introdotti e poi adottati dalla comunità scientifica – comportavano nell'identificazione stessa degli indici di dispersione spaziale.

Entrambe le categorie, sia i metodi a frequenza (come il metodo a quadrati) che quelli basati sulle distanze (come quello del vicino più prossimo), presentavano limiti concreti che sollecitavano la ricerca di procedure migliori[39]. Bisogna sottolineare comunque che, anche se i metodi a distanze riuscirono a godere di un successo maggiore nell'ambito della geografia umana, molti ricercatori consideravano i metodi tradizionali sostanzialmente inadeguati all'identificazione dei valori reali della dispersione dei sistemi di stanziamento[40].

Il problema alla base dei metodi d'analisi tradizionale è che forniscono elementi parziali o comunque non sufficientemente validi per la formulazione di comparazioni assolute tra sistemi di stanziamento. L'indice di dispersione R, ad esempio, appare fin troppo sintetico per potere essere effettivamente sfruttato nella formulazione di una qualsiasi forma di confronto efficace tra una o più forme di stanziamento. D'altra parte, i metodi classici del *point pattern analysis* non permettono di formulare elementi formali sufficienti, che in qualche modo consentano la costruzione di un qualche modello della maglia di distribuzione. È ovvio, ad esempio, che l'indice R, risultando troppo sintetico, è dunque poco attendibile come strumento per formulare una modellistica spaziale chiara e adeguata agli obiettivi dello studio delle reti di stanziamento umano[41].

6.6.1 Proprietà di secondo ordine

In precedenza è stato segnalato come la variazione di scala possa comportare un'alterazione, anche consistente, dei valori relativi agli indici tradizionali[42]. Fattore che è direttamente collegato a quella che nel campo dell'analisi spaziale è definita come "proprietà di secondo ordine".

[39] Haining 1982, p. 212.

[40] Ripley 1977, pp. 172-173: «One of the drawbacks of existing methods is their inability to test for interactions at different scales simultaneously in formal significance tests.»

[41] Haining 1982, pp. 211-212. Per un'introduzione alla modellazione dei processi per punti in geografia con metodi di secondo ordine si consiglia Getis 1983, pp. 75-76.

[42] La variazione della dispersione sulla base della scala di misurazione è un concetto introdotto prematuramente (se confrontato con i metodi classici) da Bartlett 1964, con il quale venivano gettate le basi per l'applicazione successiva dei metodi di secondo ordine.

Il concetto di secondo ordine non è immediatamente intuibile e richiede qualche stratagemma espositivo per essere reso più chiaro. Si prenda in considerazione il processo casuale indipendente (vedi paragrafo 4.9.3) ed in particolare la condizione di indipendenza. È chiaro che tale principio appare ben lontano dalla realtà materiale di una rete di stanziamento. Il problema alla base è che ogni elemento della rete di distribuzione reale finisce per condizionare i vicini[43]. Non si può non richiamare a questo punto "la prima legge della Geografia di Tobler": «Everything is related to everything else, but near things are more related than distant things.»[44]

Dato un certo territorio o area geografica, e data una rete di stanziamento al suo interno, si potrà affermare che in quest'ultima ci saranno tendenze generali che definiremo "a larga scala". Le tendenze generali vanno associate al principio di "primo ordine", e sono quelle che vengono identificate dallo studio di processi generali, come nel caso del metodo dei quadrati. Avremo poi processi che avranno un effetto solo a livello locale. Processi che finiscono per violare il principio di indipendenza implicito nelle procedure del *point pattern analysis* classico. Con il concetto di secondo ordine vengono identificati questi processi locali[45].

Per capire la distinzione tra processi di primo e secondo ordine è necessario immaginarli come più o meno coerenti con un processo soggiacente.

Come esempio della distinzione tra processi di primo e secondo ordine ne prenderemo uno tipico della geografia medica: la distribuzione delle neoplasie su una certa regione. Sarebbe ragionevole partire dalla deduzione che la distribuzione dei casi di cancro all'interno di un dato territorio segua proporzionalmente l'andamento della popolazione. Lì dove ci sono più persone ci saranno più casi; dove la densità è inferiore, meno malati. Immaginando questi casi riportati su una carta per punti, la loro distribuzione, in questo caso, rappresenta un processo di primo ordine. Cosa succede invece se all'interno di questo modello vengono introdotte diverse fabbriche responsabili dalla immissione nell'ambiente di sostanze altamente inquinanti responsabili dell'insorgere del particolare tipo di neoplasia di cui stiamo parlando? Dovremmo a questo punto immaginare una crescita rilevante dei valori, oltre quelli normali, intorno a questi punti specifici. Questo incremento a livello spaziale dei casi di malattia viene inteso come processo di secondo ordine.

Scopo dei nuovi metodi d'analisi spaziale è l'identificazione non solo del livello di dispersione o concentrazione, ma la misurazione e analisi di tali indici

[43] Haining 1982, p. 211: «Only rarely can the location of an object on the map be considered independent of the locations of other objects within the area.»

[44] Tobler 1970, p. 236.

[45] O'Sullivan, Unwin 2003, pp. 79-80.

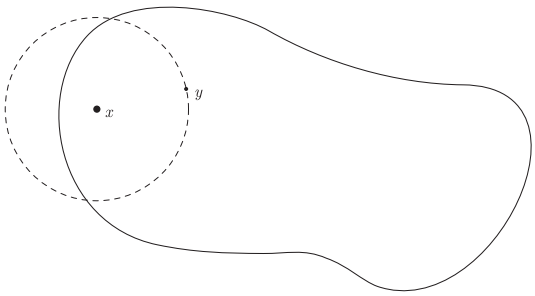


Fig. 6.15. Nella procedura di quantificazione illustrata dall'equazione 6.16 i valori di frequenza sono espressi (o convertiti) in termini di superficie, permettendo così di sottrarre la parte esterna ai confini dell'area di studio garantendo risultati più accurati per le aree di confine. Consultare Ripley 1981, p. 160; Getis 1983, p. 75.

su più scale spaziali, oltre all'identificazione di eventuali processi di secondo ordine. Da ciò il concetto di analisi multi-spettrale. Per spettro si intende proprio la variazione di scala[46].

Il metodo ad avere guadagnato maggiore prestigio è stata la cosiddetta funzione K di Ripley. Procedura che presenta alcune caratteristiche principali che occorre illustrare prima di procedere a una sua spiegazione pratica. Il metodo si fonda idealmente sulla combinazione dei metodi a distanza e metodi a frequenza. In qualche misura si rende esplicita la necessità di combinare questi due parametri per potere comprendere l'organizzazione della maglia su più scale[47]. Un altro punto importante è che la funzione K, pur non essendo complessa in termini statistici, richiede cicli di calcolo molto estesi che la relegano esclusivamente nell'ambito delle applicazioni possibili solo grazie all'ausilio del calcolatore. Il metodo qui trattato richiede il rilevamento intenso di misurazioni su più scale, impossibile in termini pratici (ma non in termini teorici) con i metodi tradizionali. In questo senso l'aspetto da considerare non è l'uso del calcolatore, bensì a quale applicativo *software* ci si debba affidare.

La funzione K, tra le altre cose, risolve diversi limiti dei metodi precedenti. Si pensi anche al problema della misurazione della dispersione in territori caratterizzati da confini complessi.

[46] Getis, Franklin 1987, pp. 473-474: «The analysis is second order because it is the variation rather than the mean of distances that is being studied».

[47] Getis 1983, p. 74: «The object is to find a cumulative distribution function based on all distances between pairs of objects».

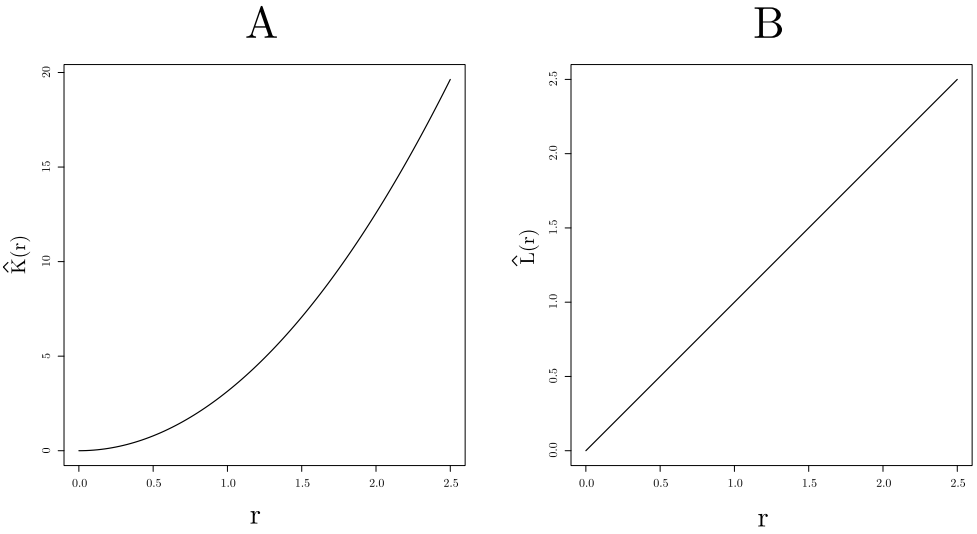


Fig. 6.16. A) Rappresentazione della curva attesa per una distribuzione di Poisson. B) Linearizzazione della medesima curva grazie all'equazione 6.18.

Il metodo è solitamente descritto da una funzione veramente poco intuibile:

$$\hat{K}(r) = \sum_{x=1}^{n} \sum_{y=1}^{n} k(x,y)/\lambda^2 A \tag{6.15}$$

dove r è il raggio di misurazione, λ rappresenta la densità attesa (n/A), A l'area del territorio sotto esame e $k(x,y)$ un operatore che definisce un'area circolare rappresentativa dei valori osservati (questo stimatore sarà discusso subito dopo)[48]. Un altro modo di rappresentare la funzione K è quella della notazione del suo stesso inventore[49]:

$$\hat{K}(r) = A\left(\sum k(x,y) \right)/N^2 \tag{6.16}$$

dove N rappresenta il totale dei punti della rete sotto analisi. Indubbiamente il punto più criptico della formula (sia in 6.15 che in 6.16) è $k(x,y)$. Questa notazione specifica è stata adottata da Ripley per indicare il computo dei numeri osservati all'interno delle aree analizzate. Tale valore rappresenta la

[48] Fotheringham, Brunsdon, Charlton 2000, pp. 150-151; Haining 1982, p. 216; O'Sullivan, Unwin 2003, pp. 93-94; Getis 1983, p. 74.

[49] Ripley 1981, pp. 158-159.

superficie di un cerchio che ha come centro x e ha una circonferenza che passa dal punto y. Il cerchio ha un'area proporzionale al numero di punti che cadono al suo interno. In altre parole, la frequenza viene rappresentata come una superficie. Il motivo di questa scelta, cioè rappresentare la frequenza all'interno del cerchio come area e non come numero di punti o densità, risponde alla necessità di dare risposta al problema degli effetti del bordo dell'area di studio sui risultati finali. Adottando una superficie si può sottrarre e scartare la parte esterna al confine. Il parametro di riferimento per il computo di $k(x, y)$ è il processo di Poisson. Il seguente esempio può servire a rendere le cose più chiare. Si immagini un territorio D con una superficie di 10.000 km^2 e una distribuzione di 600 punti. In questo caso $\lambda = 0.06$. Partendo da un punto x viene tracciato un cerchio del tipo $k(x, y)$ con un raggio di 4 km. Si contano i punti all'interno e si rileva che oltre al punto x sono presenti altri 2 vicini; dunque un totale di 3 punti. A questo punto si procede a calcolare il valore equivalente dei 3 punti in termini di superficie del cerchio. Un cerchio con un raggio di 4 km ha una superficie di $50.265km^2$. Dunque, secondo una distribuzione omogenea di Poisson, all'interno di un cerchio simile sono attesi 3.015 punti ($50.265 \times \lambda$). In questo caso 3 punti saranno una proporzione tra

$$\begin{aligned} 3 &\longrightarrow x \\ 3.0159 &\longrightarrow 50.265 \end{aligned}$$

ovvero

$$x = \frac{3 \times 50.2654}{3.01593} = 50.0004$$

che sta ad indicare un cerchio leggermente più piccolo di quello atteso[50]. Nel processo di calcolo della superficie viene tenuto conto di quanta parte della superficie del cerchio (di misurazione) cade fuori dai confini dell'area sotto esame. Nel calcolo della proporzione di $k(x, y)$ si tiene conto anche di quella misurazione.

Il processo di applicazione della funzione K prevede la misurazione reiterata dei valori di frequenza partendo dalle immediate vicinanze, con identificazione dei vicini ad intervalli (raggi) sempre maggiori. Questi raggi crescenti vanno intesi come una variazione continua di scala; e per ognuna di queste si registra la media delle superfici utilizzate (come spiegato sopra) per rappresentare le frequenze osservate. Per comprendere al meglio il metodo è utile osservare la figura 6.17. In questa si può notare una serie di cerchi concentrici a partire da

[50] Ripley 1979, p. 369.

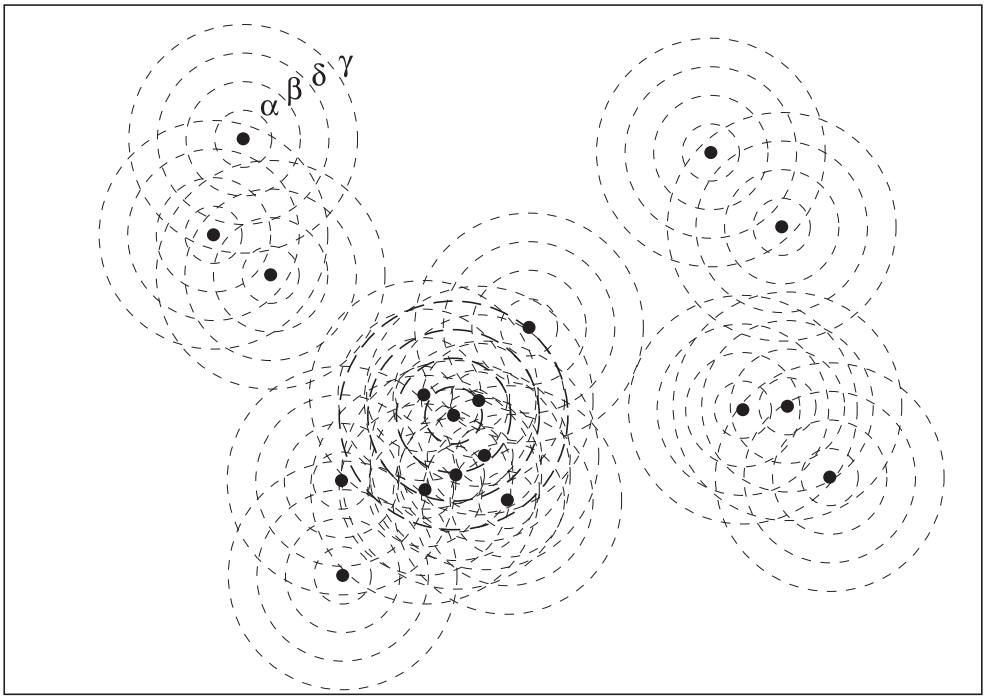


Fig. 6.17. La funzione K di Ripley si basa sulla misurazione su più scale dei valori di frequenza. Tale scopo si ottiene incrementando progressivamente ad intervalli regolari il raggio di misurazione t: α, β. . . . I valori attesi – ovvero il numero di vicino osservati – saranno equivalenti a quelli della distribuzione di Poisson.

ogni punto della maglia sotto esame. In estrema sintesi, il metodo consiste nel misurare su più distanze la densità osservata. Perciò quale sia la densità dei punti registrata a 1 km (α), a 2 km (β) e così via fino, ad esempio, a 20 km o a qualsiasi altra misura a seconda dell'estensione del territorio analizzato. In realtà il metodo prevede misurazioni ad intervalli più fitti. E, come è stato già suggerito, nel calcolo della densità si tiene conto anche dei punti generatori.

In base a queste considerazioni si può dedurre che, nel caso di una distribuzione di Poisson:

$$\hat{K}(r) = \pi \times r^2 \tag{6.17}$$

Questa ultima equazione serve per tracciare agevolmente il grafico della curva teorica o attesa (figura 6.16, A). In altre parole, come nella tradizione del *point pattern analysis*, il confronto viene realizzato con l'andamento di una

distribuzione arbitraria di tipo stocastico[51]. A questa andranno a sovrapporsi poi i valori osservati e sarà possibile fare così un raffronto tra valori osservati e attesi. Per favorire il confronto tra valori osservati e valori attesi, è suggerito di procedere alla linearizzazione della funzione grazie alla seguente formula[52]:

$$\hat{L}(r) = (\hat{K}(r)/\pi A)^{1/2} \tag{6.18}$$

dove $\hat{L}(t)$ rappresenta semplicemente $\hat{K}(t)$ linearizzata; la rappresentazione attesa di $\hat{L}(t)$ nel caso di una distribuzione di Poisson si può osservare nella figura (figura 6.16, B)

In sostanza, la funzione vuole indicare la somma, per una distanza x, di tutti i vicini registrati; in questo modo è possibile compiere un confronto con la densità generale (λ).

Le misurazioni registrate vengono successivamente confrontate con i valori attesi secondo una distribuzione di punti casuale.

6.6.2 Lo studio di *patterns* rurali con la funzione K di Ripley

Già altrove è stata proposta una distinzione marcata dei *trend* di stanziamento della rete castrense della Toscana durante i secoli centrali del medioevo. Le conclusioni alle quali si era giunti è che il territorio della regione era suddiviso geograficamente in due zone caratterizzate da *patterns* insediativi diversi: la fascia settentrionale (denominata fascia A) con una tendenza verso l'aggregazione o *clustering* e la zona meridionale della Toscana, o fascia B, con una rete dispersa[53].

Grazie a questa distinzione geografica si era proceduto al calcolo della dispersione con l'aiuto dell'indice di dispersione R, identificando per la fascia A un indice di dispersione di 0.97 e per la B di 1.11[54].

[51] I test di significatività per le distibuzioni con l'ausilio della funzione K sono illustrati in Ripley 1979. Inoltre consultare Getis, Franklin 1987, p. 474.

[52] Haining 1982, p. 216; Fotheringham, Brunsdon, Charlton 2000, p. 153.

[53] Macchi 2007, p. 91 (fig. 37) e pp. 88-89.

[54] Macchi 2007, p. 195.

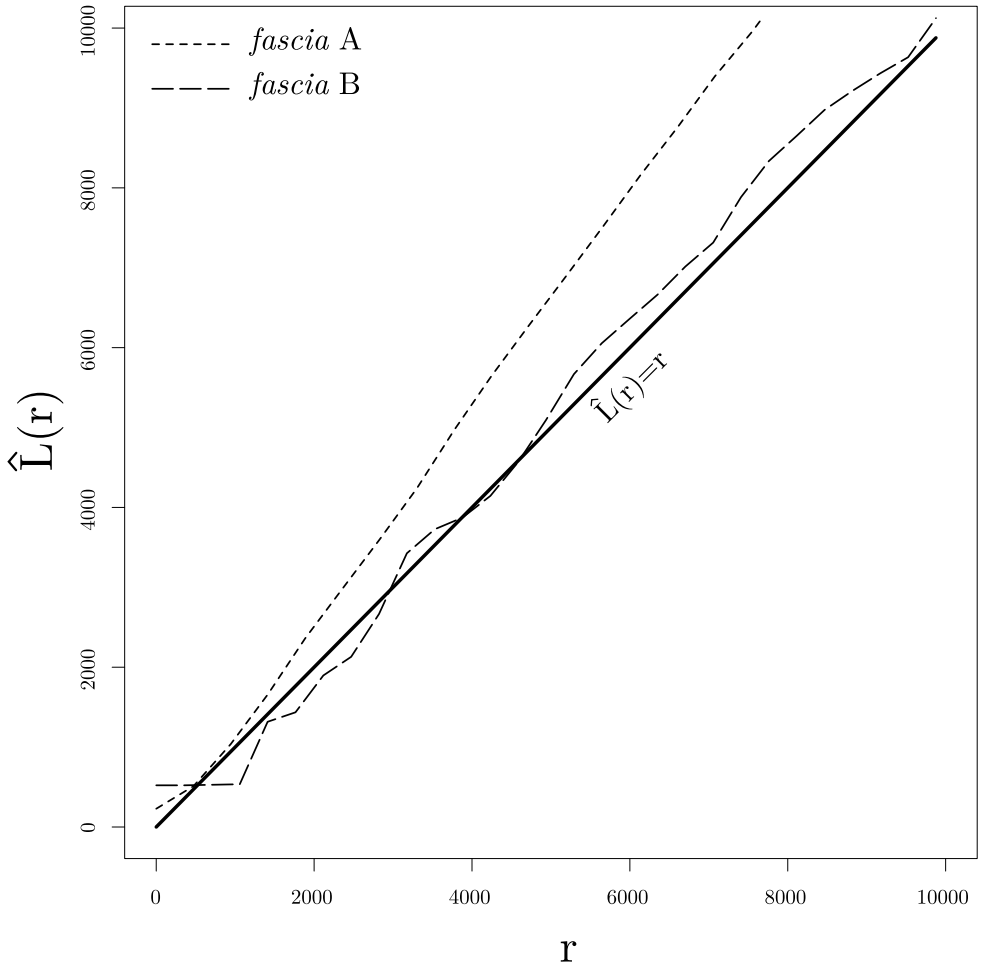


Fig. 6.18. Confronto tra i valori di $\hat{L}(t)$ per le maglie castrensi dei territori settentrionali (fascia A) e meridionali (fascia B) della Toscana relative ai secoli centrali del medioevo.

Procedendo dunque al calcolo di z con l'equazione 6.13, si ottiene $z = -1.43$; dunque un valore che permette di confutare con molta difficoltà, ovvero 92%; (un valore superiore al 95% permetterebbe invece di negare con maggior grado di certezza che la dispersione sia casuale) il fatto che, nel caso della fascia A la distribuzione di punti non sia casuale, ovvero non sia stata generata da un processo assimilabile a quello di Poisson. Nel caso del territorio meridionale, il computo della significatività, eseguito sempre con l'aiuto di 6.13, porta a $z = 2.19$, valore che permette invece di escludere, con un livello di confidenza

del 98.6%, che la distribuzione della fascia B sia stata generata da un processo casuale indipendente. In sintesi, l'analisi del vicino prossimo ci permette di individuare un probabile processo accentrato nella fascia A e un quasi certo processo disperso in quella B. L'indice di dispersione R consente in questo caso solo di distinguere i due sistemi come tendenzialmente diversi. Certamente un elemento interessante ma che poco aggiunge alla comprensione delle due strutture distributive.

Procedendo all'applicazione della funzione K di Ripley e quindi alla linearizzazione grazie al pacchetto *Spatstat* si sono ottenuti i risultati illustrati nella figura 6.18. Come si può leggere dal grafico, emergono elementi rilevanti sulla struttura delle due maglie di stanziamento. Per prima la maglia relativa alla fascia A si può notare come, a scala locale (parte sinistra delle ascisse), essa sia caratterizzata effettivamente da una struttura arbitraria conforme all'andamento del processo casuale indipendente. Da questo elemento si può spiegare l'ambiguità dell'indice di dispersione R in questa zona. La variazione di scala, però, determina un progressivo e netto incremento dell'aggregazione spaziale. La logica della struttura spaziale di questa maglia di stanziamento su scala generale indica un rilevante processo di concentrazione. Al contrario, la fascia B si dimostra caratterizzata da un'aderenza complessiva, su tutto lo spettro delle scale analizzate, ad una distribuzione casuale. Per essere più precisi, a livello locale sembra esservi una qualche forma di tendenza verso la dispersione; fattore che può in qualche misura spiegare il valore di $R = 1.11$.

Gli elementi che emergono dalla lettura di questi risultati possono così favorire una ricostruzione e una più approfondita descrizione e comprensione delle strutture spaziali.

7
Riflessioni conclusive

Non possono essere molte né tanto meno significative le conclusioni al termine di queste pagine. Ciò è dovuto principalmente al fatto che i contenuti fin qui esposti non sono che l'introduzione ai metodi quantitativi, e peraltro solo a una parte di essi, e a una indubbiamente ristretta porzione delle possibili loro applicazioni nell'ambito della geografia.

Il testo ha infatti una valenza esclusivamente introduttiva. Come è stato già sottolineato fin dalle prime righe quelli proposti sono solo gli elementi minimi ed essenziali utili allo studente che si accinga ad iniziare un percorso nell'ambito dei metodi quantitativi e normativi nello studio della geografia umana, geografia storica, storia e archeologia dei paesaggi.

Presumibilmente, per molti, discutere oggi di geografia quantitativa rappresenta un salto nel passato. Lo scopo non è quello di ritornare su questioni vecchie di decenni, nè tanto meno di sottolineare quanto certi temi siano più o meno importanti. Non si desidera spolverare una parte dello scaffale, nè ragionare sulle eventuali cause dell'insuccesso della "rivoluzione quantitativa". Certamente essa non ha saputo aprirsi a contesti applicativi ampi, chiudendosi in una sorta di corrente metodologica senza portare mai le singolarità applicative ad affrontare questioni teoriche generali. Oggi la rivoluzione quantitativa appare forse più come una sorta di "carrozzone" nel quale, dopo una fase iniziale, più geografi hanno intrapreso linee di produzione caratterizzate da un totale ermetismo matematico formulando una serie di nozioni di scarsa o nulla applicabilità nell'ambito della geografia dell'uomo. In contrapposizione a queste complicate elaborazioni, ritengo sia da preferire una statistica descrittiva elementare, in grado di porre al centro della nostra attenzione la capacità di scambiare idee e concetti sullo spazio.

Nel compilare questo testo ho costantemente cercato di mettere l'accento su due aspetti che considero fondamentali per qualsiasi geografo: lo spazio e i suoi modelli. Come già accennato, non avrebbe molta logica mettere i modelli locazionali insieme alle procedure quantitative; i primi di stampo

deterministico-funzionali, le altre di stampo probabilistico. Quello che unisce le due branche è lo spazio: minimo comune denominatore di queste due tradizioni della geografia. Entrambe finiscono per produrre modelli, che sono, per propria natura, di tipo spaziale. Cioè schemi che, attraverso processi di semplificazione, ci aiutano a comprendere meglio la strutturazione di quello che possiamo intuire come spazio sociale.

Fra le diverse categorie ed esperienze cognitive del geografo vi è quella dello spazio. Categoria spesso omessa, per la sua essenza minimale intrecciata ad una complessità e astrazione che la caratterizzano e la rendono fugace. La complessità dello spazio – la sua natura multidimensionale o scalare – viene messa in evidenza ogni qual volta il geografo, o lo studioso del territorio, procede ad un cambio di scala. I fenomeni sociali – ma anche quelli fisici –, in conformità con la logica della geometria frattale, cambiano a seconda della distanza (la scala) da cui si guardano.

È interessante notare come, sia per quanto riguarda l'area dei modelli geografico-economici classici, che per quanto riguarda il settore dell'analisi spaziale, tutto sembra acquisire nuova vitalità quando la scala geografica entra a far parte dello schema cognitivo. Sia Lösch, con la sua sovrapposizione di territori variabili, che la funzione K di Ripley dimostrano come un ragionamento più profondo sulla scala geografica sia necessario per la comprensione delle forme di occupazione sociale dello spazio.

7.1 Ostacoli sul percorso

Lo stato ancora embrionale dell'informatica applicata nel momento della massima auge della rivoluzione quantitativa ha sicuramente costituito un limite per quest'ultima. I pionieri della geografia quantitativa hanno solo in rari casi potuto contare su macchine di calcolo digitale e naturalmente solo in modo episodico. La rarità del calcolo digitale non costituiva l'unico limite alla cristallizzazione di una geografia quantitativa.

Oltre ai metodi e ai calcolatori, occorreva soprattutto la costruzione scrupolosa e critica delle banche dati sulle quali applicare i metodi. Processo che si verifica nei contesti disciplinari prettamente umanistici solo attorno alla metà degli anni '90.

Si vuole perciò mettere in evidenza come la rivoluzione attualmente in atto – ovvero la costruzione di grandi archivi digitali di dati spaziali – porterà in un futuro non molto lontano ad una fase successiva, cioè quella della loro analisi critica e formale, determinando la necessità di riscoprire quella parte della tradizione metodologica. Ultimato il processo di costruzione di una banca dati georiferita, si procederà sempre a domandarsi: "e adesso cosa bisogna fare?"

La risposta sarà sempre una sola: "analizzarla spazialmente nei migliori dei modi", ovvero cercare di rendere i risultati chiari e comprensibili a tutta la comunità scientifica. La diffusione delle tecnologie GIS nell'ambito delle scienze umane non è andata di pari passo con quella degli approcci analitici. Questo aspetto verrà esaminato subito dopo.

Ma questo naturalmente non è l'unico aspetto degno di nota. Un altro settore colpito dall'analfabetismo quantitativo è la cartografia scientifica, cioè quella prodotta nell'ambito della ricerca. Cartografare la conoscenza – o ancora di più la nuova conoscenza derivante dalle ricerche in corso – richiede un dominio della semantica cartografica; che è, secondo i suoi propri principi intrinseci, un mezzo di comunicazione di tipo formale «Communication by thematic maps is a highly complex, speculative venture. It is easy to "draw" a map and declare the endeavor completed. This, however, is not the goal of the cartographer; he attempts to transmit a message successfully. Cartography is not an exercise in draftsmanship, but a scientific enterprise founded [...]»[1] Per essere efficaci, il simbolo, la tonalità e il colore devono saper tradursi nella conoscenza di fondo che si vuole trasmettere. All'interno della cartografia scientifica poco spazio resta per interpretazioni soggettive. Quello che qui si vuole mettere in evidenza è che se, nell'ambito della ricerca, si desidera creare carte con un qualche significato critico, occorre conoscere i principi fondamentali dell'analisi dei dati. Principi che naturalmente sono non solo la chiave di codifica, ma anche di decodifica. I sistemi GIS, per quanto abbiano automatizzato certi processi, non potranno supplire alle necessità espositive del ricercatore. Semplicemente, come un *editor* di testi non potrà mai scrivere un romanzo per conto dell'utente, un sistema GIS non potrà mai, in modo automatico, compilare la carta più adatta.

Pur nel progresso tecnologico degli ultimi vent'anni si è assistito parallelamente ad un impoverimento del materiale cartografico delle discipline umane, e non solo: in certi settori si è giunti ad una sua progressiva scomparsa. Forse tutto questo può essere imputato ad un'evoluzione ciclica delle discipline, ma credo invece che parte di questa "povertà cartografica" sia da attribuire proprio ad un certo tipo di informatizzazione che ha caratterizzato certi settori disciplinari. A questo tipo di banalizzazione informatica appartiene anche l'uso improprio che diverse scienze umane fanno di applicativi GIS.

È interessante notare come oggi, a conferma di ciò cui si accennava sopra, il materiale cartografico di qualità viene prodotto esclusivamente da quelle istituzioni che già in passato, prima dell'introduzione dei GIS, dominavano le tecniche cartografiche.

[1] Dent 1972, p. 93.

7.2 I *patterns*

A più riprese è stato evidenziato come la statistica in campo geografico non abbia di per sè un potere tautologico. L'identificazione dei *patterns* spaziali ha poco significato, soprattutto se gli stessi non sono compresi. Questo per due motivi essenziali. Da una parte, chi non è interessato ai *patterns* non è detto che sia effettivamente in grado di identificarli. L'altro aspetto è che, anche se si può credere che nella realtà fisica e sociale esistano delle trame o tendenze generali, allora è necessario attribuire a queste un valore esplicativo dei fenomeni studiati. In altre parole, non ha molto significato misurare per misurare: il tutto si risolverebbe in un descrittivismo inutile.

Va detto perciò che l'identificazione dei *patterns* di per sè non costituisce un traguardo, ma una tappa del processo d'indagine. Chiariamo subito che l'analisi dei dati spaziali può permettere al ricercatore di identificare dei *patterns*, ma mai di interpretarli, nè tanto meno di spiegarli. Questo spetta naturalmente alla sua capacità intuitiva e al *background* disciplinare.

Sbaglia colui che considera che la geografia quantitativa possa fornire qualcosa più di questo. Ma il fatto che l'analisi spaziale permetta di identificare i *patterns*, ma non di spiegarli, non ci autorizza a considerare questa branca una sorta di inutile accessorio disciplinare. Perché si tratta di un settore che può compiere operazioni di lettura del paesaggio umano in modo preciso, garantendo inoltre la condivisione aperta e chiara delle proprie interpretazioni. L'analisi spaziale ci permette di giocare a carte scoperte.

Si giunge alla considerazione che i patterns distributivi – di qualsiasi natura – passano attraverso un approccio di tipo analitico che poco spazio lascia alla lettura qualitativa di tipo soggettivo.

Ed è naturalmente il punto che spesso si tende a sottovalutare. L'analisi spaziale non fornisce spiegazioni, permette però di formulare apparati di comunicazione indispensabili. Vale a dire che l'approccio formale, a differenza di un approccio qualitativo, presenta il vantaggio di potere convogliare le proprie osservazioni (le misurazioni) in modo chiaro e preciso.

7.3 GIS e geografia quantitativa

La geografia quantitativa ha avuto un contributo straordinario grazie all'introduzione del calcolatore. Per molti versi, tale evento potrebbe essere paragonabile a quello dell'avvento del microscopio per la biologia o del telescopio per l'astronomia. Non vi è alcun dubbio che la possibilità di automatizzare processi di computo abbia contribuito ad un rilancio e potenziamento delle attività di analisi. Il calcolatore elettronico non è che, come d'altronde suggerisce

il nome, una macchina per fare i conti. Si fa fatica a comprendere che dietro le immagini a colori, i suoni e le animazioni 3D non ci sia altro che un'enorme massa di numeri. Se scendiamo nel particolare dell'architettura del calcolatore comprendiamo che esso non è altro che l'equivalente di un abaco grandissimo.

Ancora oggi è molto diffusa la convinzione secondo la quale l'analisi spaziale e la geografia quantitativa traggono un enorme vantaggio dalla diffusione dei sistemi GIS. In realtà non è assolutamente necessario disporre di un *software* GIS per compiere delle analisi spaziali. La maggior parte delle operazioni di analisi (sia in ambiente *raster* che vettoriale) e quantificazione geografica possono essere svolte benissimo con applicativi di statistica e geostatistica; addirittura certi tipi d'analisi si possono svolgere solo con questi ultimi.

Ancora più grave è l'idea secondo la quale esiste una corrispondenza tra analisi spaziale e GIS. È vero che l'analisi spaziale non è sinonimo di GIS. Una delle cause di questo equivoco può essere rintracciata nella definizione stessa di GIS. Comunemente la definizione più diffusa per questo tipo di applicazione è quella di «sottosistema informatico utile all'archiviazione, gestione e analisi di dati».

Il fatto che un sistema GIS sia effettivamente in grado di compiere alcuni tipi di analisi su dati spaziali non implica che lo stesso pacchetto *software* sia effettivamente la categoria specifica per l'esecuzione di analisi spaziali. Questo sarebbe come confondere la calcolatrice tascabile con la matematica o, quel che è peggio, colui che possiede una calcolatrice con un matematico.

D'altra parte si può notare come la diffusione dei sistemi GIS abbia in qualche misura limitato e ristretto l'interesse delle diverse comunità scientifiche ad un approccio geografico quantitativo. Paradossalmente, il grande sviluppo e la conseguente diffusione che ebbero i GIS a partire dai primi anni '80 nel campo della ricerca scientifica più avanzata, e, dagli inizi degli anni '90, l'introduzione del *desktop* GIS, hanno contribuito ad attenuare gradualmente l'interesse di cui la geografia quantitativa e gli studi spaziali in genere avevano goduto.

In questo senso, il passo in avanti più importante è stato l'introduzione degli strumenti di analisi basati sul modello *raster*. Non casualmente Fischer, il padre del *Laboratory for Computer Graphics and Spatial Analysis* e pioniere del GIS nella sua forma a noi nota, iniziò la sua opera di sviluppo e di eredità per il *software* commerciale successivo con SYMAP, ovvero un sottosistema in grado di gestire e analizzare carte di tipo *raster*. Ancora oggi le cose non sono cambiate, dato che sono davvero limitate le operazione di analisi che possono essere portate a termine in ambiente vettoriale.

Un aspetto cruciale del problema che questa breve guida ha cercato di affrontare riguarda l'uso che si può fare degli strumenti di analisi spaziale presenti all'interno dei *software* GIS. Questi programmi sono muniti di una

quantità più o meno estesa, a seconda dei casi, di strumenti analitici specifici. La loro esistenza non costituisce però garanzia di alcunché; né tanto meno esprimerebbe il necessario rigore scientifico di qualsiasi procedura analitica. Si pensi, ad esempio, al caso dell'utilizzo di questi comandi da parte di un utente privo però di nozioni e conoscenze sui metodi geografico-quantitativi. Egli arriverà sicuramente a delle assunzioni del tutto errate, ad un utilizzo (o applicazione) impropria e ad una valutazione sbagliata dei risultati. Pur disponendo degli strumenti analitici più sofisticati, il ricercatore deve possedere le nozioni minime di base per poter procedere ad analisi formali in modo critico e rigoroso. Disporre di una calcolatrice rappresenta un grosso vantaggio, ma occorre sapere quali operazioni si possono compiere e soprattutto come portarle a termine.

Una delle caratteristiche degli studi spaziali è anche la continua evoluzione dei protocolli metodologici. Un metodo d'analisi può subire innumerevoli cambiamenti in ogni parte della sua sequenza d'applicazione. Sarebbe poco coerente pensare che gli strumenti analitici presenti all'interno dei *software* GIS siano sufficienti a garantire la loro applicabilità ad ogni sorta di problema di ricerca. In poche parole, come è possibile ricorrere a strumenti fissi per risolvere problemi la cui morfologia muta costantemente?

D'altro canto, bisogna prendere in considerazione un fattore non meno importante. Una delle caratteristiche principali dei sistemi GIS è la costruzione di rigidi modelli dati in modo che sia garantita una corretta e adeguata archiviazione e gestione dei dati. Il modello dei dati diviene così un ambiente sterile per compiere delle analisi spaziali. In effetti è il *feedback* costante a caratterizzare un processo analitico.

Solitamente le operazioni di analisi spaziale vengono realizzate all'interno di processi disorganici, dove difficilmente è possibile stabilire in anticipo quale sarà la successiva mossa tesa all'identificazione di possibili correlazioni. Questa procedura disgregata poco si addice alla natura formale dei modelli dei dati. D'altronde, se le analisi spaziali vengono compiute con il fine di comprendere meglio le caratteristiche dell'oggetto di studio, come sarebbe possibile definirne le caratteristiche in anticipo come la costruzione di un modello dei dati richiederebbe?

7.4 La complessità

In un futuro – auspicabilmente non molto lontano – il geografo, più che con classificazioni statistiche o analisi di dispersione spaziale, avrà a che fare con strumenti di natura euristica tesi a facilitare il processo di assimilazione e comprensione della realtà. Mi riferisco in particolare ad un settore composto

prevalentemente dalla scienza dei frattali, la scienza dei simulatori e l'intelligenza artificiale; quest'ultima scomposta nei campi delle reti neurali artificiali e degli algoritmi genetici.

Un po' per l'evoluzione spontanea delle problematiche di studio, e un po' per il trascinamento provocato dalla scia di altre discipline che da tempo hanno cominciato il proprio percorso verso la complessità, la geografia ha già iniziato ad adottare questi metodi come parte del proprio corpo metodologico. Si potrebbe aggiungere che, in taluni casi, queste applicazioni nascono – almeno in parte – nell'ambito della geografia, come nel caso dei frattali e della famosa domanda *How long is the coast of Britain?*

Sotto l'ottica della nuova e ormai imminente rivoluzione (in parte già in atto) delle nuove domande che la geografia umana e tutte le scienze collegate andranno a porsi, gran parte delle nozioni contenute in questo testo non appariranno molto utili ai fini della costruzione dei nuovi strumenti concettuali. Non credo sia molto opportuno, né tanto meno produttivo, vivere la propria attività scientifica seguendo le mode disciplinari, dimenticando troppo spesso quello che di buono ogni generazione porta con sé; scartando questo o quello perché appartiene a quella scuola, o, peggio ancora, a un determinato periodo. Nel guardare con interesse agli orizzonti delle scienze della complessità non si possono trascurare la natura e il significato della geografia quantitativa. Se non altro perché questo rappresenta in ogni caso l'ambito entro il quale nasceranno o verranno prodotte le informazioni e le osservazioni utili alle nuove forme di conoscenza.

Riferimenti bibliografici

Aase T.H., 1994, *Representations of space in geography and anthropology*, «Geografiska Annaler», vol. 76, n. 1, pp. 51-58.

Aamodt G., 2000, *A statistical model to analyse spatio-temporal data*, in Fotheringham, Wegener 2000, pp. 235-244.

Abel D.J., Kilby P.J., Davis J.R., 1994, *The system integration problem*, «IJGIS», VIII, n. 1, pp. 1-12.

Abrahart R.J., Kirkby M.J, McMahon M.L., 1994, *A combined geographical information system and large scale distributed process model*, in *Procedings of the GIS Research UK Conference*, Fisher P. (a cura di), Association of Geographical Information, Leicester, pp. 67-75.

Ackermann R.J., 1966, *Confirmatory models of theories*, «The British Journal for the Philosophy of Science», vol. 16, n. 64, pp. 312-326.

Appleby P.H., 1954, *Managing complexity*, «Ethics», vol. 64, n. 2, pp. 79-99.

Auerbach F., 1913, *Das Gesetz der Bevlkerungskoncentration*, «Petermanns Geographische Mitteilungen», vol. 59, pp. 74-76.

Ballard E.G., 1949, *The paradox of measurement*, «Philosophy of science», vol. 16, n. 2, pp. 134-136.

Bartlett M.S., 1964, *The spectral analysis of two-dimensional point processes*, «Biometrika», vol. 51, n. 3-4, pp. 299-311.

Bellone E., 2000, *I corpi e le cose: un modello naturalistico della conoscenza*, Bruno Mondadori, Milano.

Berry B.J.L., Garrison W.L., 1958, *Alternate explanations of urban rank-size relationships*, «Annals of the Association of American Geographers», vol. 48, n. 1, pp. 83-91.

Berry B.J.L., Harris C.D., 1970, *Walter Christaller: an appreciation*, «Geographical Review», vol. 60, n. 1, pp. 116-119.

Bloch M., 1969, *Apologia della storia*, Einaudi, Torino.

Bohrnstedt G.W., Knoke D., 1998, *Statistica per le scienze sociali*, Il Mulino, Bologna.

Boots B.N., 1980, *Weighting Thiessen polygons*, in «Economic Geography», vol. 56, n. 3, pp. 248-259.

Boots B.N., 1987, *Modifying Thiessen polygons*, in «The Canadian Geographer», vol. 31, n. 2, pp. 160-169.

Braudel F., 1969, *Histoire et sciences sociales. La longue durée*, *Ecrits sur l'histoire*, Flammarion, Paris, pp. 41-83.

Braudel F., 1976, *Civiltà e imperi del mediterraneo nell'età di Filippo II*, Einaudi, Torino.

Braudel F., 1982, *La storia e le altre scienze sociali: antologia di saggi delle Annales E.S.C.*, Laterza, Roma.

Buttimer A., 1969, *Social space in interdisciplinary perspective*, «Geographical Review», vol. 59, n. 3, pp. 417-426.

Burke P., 1992, *Una rivoluzione storiografica : la scuola delle "Annales" (1929-1989)*, Laterza, Roma.

Byerly H.C., Lazara V.A., 1973, *Realist foundations of measurement*, «Philosophy of science», vol. 40, n. 1, pp. 10-28.

Camassa G., De Guio A., Veronese F. (a cura di), 2000, *Paesaggi di potere, problemi e prospettive*, Atti del Seminario (Udine, 16-17 Maggio 1996), Qasar, Roma.

Carr D., 2001, *Place and time: on the interplay of historical points of view*, «History and Theory», vol. 40, n. 4, pp. 153-167.

Carter H., 1975, *La geografia urbana: teoria e metodi*, Zanichelli, Bologna.

Cartwright N., Shomar T., Suárez M., 1995, *The tool box of science*, «Poznan Studies in the Philosophy of the Sciences and the Humanities», vol. 44, pp. 137-149.

CHRISTALLER W., 1933, *Die zentralen Orte in Süddeutschland*, Gustav Fischer, Jena.

CHRISTALLER W., 1980, *Le località centrali della Germania Meridionale. Una indagine geografica sulla regolarità della distribuzione dello sviluppo degli insediamenti con funzioni urbane,* Franco Angeli, Milano.

CHURCHILL SEMPLE E., 1908, *Geographical location as a factor in history*, «Bulletin of the American Geographical Society», vol. 4, n. 2, pp. 65-81.

CHORLEY R.J., 1965, HAGGETT P., *Models in geography*, Methuen, London.

CHORLEY R.J., HAGGETT P., 1969,*Network analysis in geography*, Arnold, London.

CLARK P.J., EVANS F.C., 1954, *Distance to nearest neighbor as a measure of spatial relationships in populations*, «Ecology», n. 35, 1954, pp. 445-453.

CLARKE D.L., 1972, *Models in Archaeology*, Methuen, London.

CLARKE D.L., 1977, *Spatial Archaeology*, Academic Press, New York.

CLARKE D.L., 1978, *Analytical Archeology*, Columbia University Press, New York.

CLAVAL P., 1980, *L'evoluzione storica della geografia umana*, Franco Angeli, Milano.

CLAVAL P., 1983, *Elementi di geografia umana*, UNICOPLI, Milano.

CLAVAL P., 1996, *Introduzione alla geografia regionale*, Zanichelli, Bologna.

CLEMENCE G.M., 1949, *Relativity effects in planetary motion*, «Proceedings of the American Philosophical Society», vol. 93, n. 7, pp. 532-534.

CONTI E., 1965, *La formazione della struttura agraria moderna nel contado fiorentino, Le campagne nell'età precomunale*, Istituto storico italiano per il Medio Evo, Roma.

CONWAY F., 1979, *The rank size rule in perspective*, «The Mathematical Gazette», vol. 63, n. 423, pp. 30-36.

COPPOCK J.T., JOHNSON J.H., 1962, *Measurement in human geography*, «Economic Geography», vol. 38, n. 2, pp. 130-137.

COUCLELIS H., GALE N., 1986, *Space and spaces*, «Geografiska Annaler», vol. 68, n. 1, pp. 1-12.

Coulson M.R.C., 1978, *"Potential for variation": a concept for measuring the significance of variations in size and shape of areal units*, «Geografiska Annaler», vol. 60, n. 1, pp. 48-64.

D'Ercole V., 2000, *I "paesaggi di potere" dell'Abruzzo protostorico*, in Camassa, De Guio, Veronese (a cura di) 2000, pp. 121-152.

Dacey M.F., 1958, *Analysis of map distributions by nearest neighbour methods*, Discussion Paper No. 1, Department of Geography - University of Wahington, Seatle.

Dacey M.F., 1960, *A note on the derivation of nearest neighbor distances*, «Journal of Regional Science», vol. 2, n. 2, pp. 81-87.

Dacey M.F., 1965, *The geometry of central place theory*, «Geografiska Annaler», Series B, vol. 47, n. 2, pp. 111-124.

Darwin Ch., 1888, *Movements and habits of climbing plants*, Appleton, New York.

David F.N., Moore P.G., 1954, *Notes on contagious distributions in plant populations*, «Annals of Botany», vol. 18, pp. 48-53.

Dematteis G., 1982, *Prefazione*, in Vagaggini, pp. XI-XII.

Dematteis G., 1994, *Le metafore della Terra*, Feltrinelli, Milano.

Dent B.D., 1972, *Visual organization and thematic map communication*, «Annals of the Association of American Geographers», vol. 62, n. 1, pp. 79-93.

Diggle P., 1983, *The statistical analysis of point patterns*, Accademic Press, London.

Dingle H., 1950, *A theory of measurement*, «The British Journal for the Philosophy of Science», vol. 1, n. 1, pp. 5-26.

Douglas J.B., 1975, *Clustering and aggregation*, «Sankhya», Series B, vol. 37, pp. 398-417.

Dunn E.S., 1955, *Review: The Economics of Location*, «Journal of Farm Economics», vol. 37, n. 2, pp. 376-377.

Eaton B.C., Lipsey R.G., 1976, *The non-uniqueness of equilibrium in the loschian location model*, «The American Economic Review», vol. 66, n. 1, pp. 77-93.

EELLS W.C., SVIATLOVSKY E.E., 1937, *The centrographical method and regional analysis*, in «Geographical Review», vol. 27, n. 2., pp. 240-254.

ELDREDGE N., 2000, *Le trame dell'evoluzione*, Raffaello Cortina, Milano.

EVANS I.S., 1977, *The selection of class intervals*, «Transactions of the Institute of British Geographers», New Series, vol. 2, n. 1, pp. 98-124.

FETTER F.A., 1930, *Review: Alfred Weber's theory of the location of industries*, «The Journal of Political Economy», vol. 38, n. 2, pp. 232-234.

FLANNERY K., 1976, *The early messoamerican village*, Academic Press, New York.

FLANNERY K., WINTER M., 1976, *Analysing households activities* in FLANNERY 1976, pp. 34-47.

FOTHERINGHAM A.S., 2000, *GIS-based Spatial modelling: a step forward or a step backwards?*, in FOTHERINGHAM, WEGENER 2000, pp. 21-30.

FOTHERINGHAM A.S., WEGENER M., 2000, *Spatial models and GIS. New potentials and new models*, Taylor & Francis, London.

FOTHERINGHAM A.S., BRUNSDON CH., CHARLTON M., 2000, *Quantitative geography: perspective on spatial data analysis*, Sage, London.

FOUCAULT M., MISKOWIEC J., 1986, *Of other spaces*, in «Diacritics», vol. 16, n. 1., pp. 22-27.

FOX KELLER E., 2000, *Models of and models for: theory and practice in contemporary biology philosophy of science*, vol. 67, n. 3, pp. S72-S86.

FREEDMAN D., PISANI R., PURVES R., 1998, *Statistica*, McGraw-Hill, Milano.

FULTON M., HOCH L.C., 1959, *Transportation factors affecting locational decisions*, «Economic Geography», vol. 35, n. 1, pp. 51-59.

GALTON A., 2000, *Qualitative spatial change*, Oxford University Press, Oxford.

GATRELL A.C., BAILEY T.C., DIGGLE P.J., ROWLINGSON B.S., 1996, *Spatial point pattern analysis and its application in geographical epidemiology*, «Transactions of the Institute of British Geographers», vol. 21, n. 1, pp. 256-274.

GARRISON W.L., 1959, *Spatial structure of the economy: II*, «Annals of the Association of American Geographers», vol. 49, n. 4, pp. 471-482.

GETIS A., 1977, *On the use of the term "random" in spatial analysis*, «The Professional Geographer», vol. 29, n. 1, pp. 59-61.

GETIS A., 1983, *Second-order analysis of point patterns: the case of Chicago as a multi-center urban region*, «Professional Geographer», vol. 35, n. 1, pp. 73-80.

GETIS A., FRANKLIN J., 1987, *Second-order neighborhood analysis of mapped point patterns*, «Ecology», vol. 68, n. 3, pp. 473-477.

GIUSTI M., GUIDI P., 1942, *Rationes Decimarum Italiae. Tuscia II. Le decime degli anni 1295-1304*, Biblioteca Apostolica Vaticana, Città del Vaticano.

GOULD P., 1967, *Review: Locational analysis in human geography by Peter Haggett*, «Geographical Review», vol. 57, n. 2, pp. 292-294.

GREPPI C., 1996a, *Cities and countryside: empirical evidences*, in DIAMANTINI C., PATASSINI D. (a cura di), *Urban Ethiopia: evidences of the 1980s*, Istituto Universitario di Architettura, Venezia.

GREPPI C., 1996b, *Population and migration*, in DIAMANTINI C., PATASSINI D. (a cura di), *Urban Ethiopia: evidences of the 1980s*, Istituto Universitario di Architettura, Venezia.

GREPPI C., 2003, *I paesaggi di un comune senese a due secoli di distanza: il caso di Murlo, 1825-1995*, in «Trame nello spazio», n. 1, All'Insegna del Giglio, Firenze, pp. 79-88.

GREPPI C., PARDI F., 2002, *Natura e storia: i paesaggi senesi fra continuità e mutamento*, «Urbanistica Quaderni», XXXVI, pp. 102-115.

GRIFFIN E., 1973, *Testing the Von Thunen theory in Uruguay*, «Geographical Review», vol. 63, n. 4, pp. 500-516.

GROTEWOLD A., 1959, *Von Thunen in Retrospect*, «Economic Geography», vol. 35, n. 4, pp. 346-355.

GUELKE L., 1997, *The relations between geography and history reconsidered*, «History and Theory», vol. 36, n. 2, pp. 216-234.

HAGGETT P., 1965, *Locational analysis in human geography*, Arnold, London.

HAGGETT P., 1967, *Review: Three pioneers in locational theory*, «The Geographical Journal», vol. 133, n. 3, pp. 357-359.

HAGGETT P., 1970, *Obituary: Walter Christaller*, «The Geographical Journal», vol. 136, n. 3, p. 500.

HAGGETT P., 1976, *Análisis locacional en la geografía humana*, Gustavo Gili, Barcelona.

HAGGETT P., 1997, *Geografia, una sintesi moderna*, Zanichelli, Bologna.

HAGGETT P., 1997, *L'arte del geografo*, Zanichelli, Bologna.

HAGGETT P., CHORLEY R.J., 1972, *Network analysis in geography*, Arnold, London.

HAGGETT P., CLIFF A., FREY A., 1977, *Locational analysis in human geography*, Arnold, London.

HAINES-YOUNG R., GREEN D.R., COUSINS S. (a cura di), 1993, *Landscape ecology and geographic information systems*, Taylor & Francis, London.

HAINING R., 1982, *Describing and modelling rural settlement maps*, «Annals of the Association of American Geographers», vol. 72, n. 2, pp. 211-223.

HALL J.R., 1992, *Where history and sociology meet: forms of discourse and sociohistorical inquiry*, in «Sociological Theory», vol. 10, n.2, pp. 164-193.

HAMMOND R., MCCULLAGH P., 1975, *Quantitative techniques in geography*, Clarendon Press, Oxford.

HARTSHORNE R., 1939, *The nature of geography: A critical survey of current thought in the light of the past*, «Annals of the Association of American Geographers», vol. 29, n. 3, pp. 173-412.

HARTSHORNE R., 1954, *Comment on "Exceptionalism in geography"*, «Annals of the Association of American Geographers», vol. 44, n. 1, pp. 108-109.

HARTSHORNE R., 1955, *"Exceptionalism in geography" re-examined*, «Annals of the Association of American Geographers», vol. 45, n. 3, pp. 205-244.

HARVEY D.W., 1966, *Geographical processes and the analysis of point patterns: testing models of diffusion by quadrat sampling*, «Transactions of the Institute of British Geographers», n. 40, pp. 81-95.

HARVEY D., 1978a, *Il linguaggio della forma spaziale,* in VAGAGGINI 1978, pp. 28-78.

HARVEY D., 1978b, *Alcune controversie metodologiche in geografia*, in VAGAGGINI 1978, pp. 21-27.

Hempel C.G., 1943, *A purely syntactical definition of confirmation*, «Journal of Symbolic Logic», vol. 8, n. 4, pp. 122-143.

Hempel C.G., 1946, *A Note on the parodoxes of confirmation*, «Mind», New Series, vol. 55, n. 217, pp. 79-82.

Hodder I., 1972, *Locational Models and the Study of Romano-British Settlement*, in Clark 1972, pp. 887-909.

Hodder I., 1982, *Symbols in action: ethnoarchaeological studies of material culture*, Cambridge University Press, Cambridge.

Hodder I., 1984, *Disertación de Ian Hodder*, in *Archeología Espacial, Coloquio sobre distribución y relaciones entre los asentamientos* (Teruel, 27-29 de Septiembre), Universidad de Zaragoza, Teruel.

Hodder I., Orton C., 1976, *Spatial analysis in archaeology*, Cambridge University Press, Cambridge.

Holmes J., 1967, *Problems in location sampling*, «Annals of the Association of American Geographers», vol. 57, n. 4, pp. 757-780.

Horvath R.J., 1969, *Von Thunen's isolated state and the area around Addis Ababa, Ethiopia*, «Annals of the Association of American Geographers», vol. 59, n. 2, pp. 308-323.

Hottes R., 1983, *Walter Christaller*, «Annals of the Association of American Geographers», vol. 73, n. 1, pp. 51-54.

Hsin-Ping Ch., 2004, *Path-dependent processes and the emergence of the rank size rule*, «The Annals of Regional Science», n. 38, pp. 433-449.

Hudson J.C., 1969, *A location theory for rural settlement*, «Annals of the Association of American Geographers», vol. 59, n. 2, pp. 365-381.

Johnson G.A., 1980, *Rank size convexity and system integration: a view from archaeology*, «Economic geography», vol. 56, n. 3, pp. 234-247.

Johnston R.J., Gregory D., Smith D., 1986, *The dictionary of human geography*, Blackwell Reference, Oxford.

Kariel H.G., 1970, *Analysis of the Alberta settlement pattern for 1961 and 1966 by nearest neighbor analysis*, «Geografiska Annaler», vol. 52, n. 2, pp. 124-130.

KENYON J.B., 1967, *On the relationship between central function and size of place*, «Annals of the Association of American Geographers», vol. 57, n. 4, pp. 736-750.

KEUNING J., 1955, *The history of geographical map projections until 1600*, «Imago Mundi», vol. 12, pp. 1-24.

KOLARS J., 1967, *Locational analysis in human geography by Peter Haggett*, «Economic Geography», vol. 43, n. 3, pp. 276-277.

KOPEC R.J., 1963, *An alternative method for the construction of Thiessen polygons*, «The Professional Geographer», vol. 15, n. 5, pp. 24-26.

KUHN T.S., 1962, *The structure of scientific revolutions*, University of Chicago Press, Chicago.

KUHN T.S., 1969, *La struttura delle rivoluzioni scientifiche: come mutano le idee nella scienza*, Einaudi, Torino.

LAKATOS I., 1995, *Sull'orlo della scienza: pro e contro il metodo*, Cortina Raffaello, Milano.

LANDINI P., MASSIMI G., 1986, *Introduzione all'edizione italiana,* in UNWIN 1986, pp. 13-46.

LEVINE D., KREHBIEL T., BERENSON M., 2002, *Statistica*, Apogeo, Milano.

LEY D., 1977, *Social geography and the taken-for-granted world*, «Transactions of the Institute of British Geographers», New Series, vol. 2, n. 4, pp. 498-512.

LLOYD P.E. DICKEN P., 1986, *Spazio e localizzazione: un'interpretazione geografica dell'economia*, Franco Angeli, Milano.

LÖSCH A., 1937, *Population cycles as a cause of business cycles*, «The Quarterly Journal of Economics», vol. 51, n. 4, pp. 649-662.

LÖSCH A., 1938, *The nature of economic regions*, «Southern Economic Journal», vol. 5, n. 1, pp. 71-78.

LÖSCH A., 1940, *Die räumliche ordnung der Wirtschaft*, Gustav Fischer, Jena.

LÖSCH A., 1959, *The economics of location*, Yale University Press, New Haven.

LONGLEY P., 2000, *Spatial analysis in the new Millenium*, «Annals of the Association of American Geographers», vol. 90, n. 1, pp. 157-165.

Macchi G., 2001, *Modelli matematici per la ricostruzione dei paesaggi storici*, «Archeologia e Calcolatori», n. XII, pp. 143-165.

Macchi G., 2003, *Introduzione all'analisi spaziale*, in «Trame nello spazio», n. 1, All'Insegna del Giglio, Firenze, pp. 11-25.

Macchi G., 2004, *L'insediamento castrense nella Toscana (secoli X-XIII). Analisi spaziale e studi quantitativi*, tesi di Dottorato di Ricerca (XVI ciclo), Università degli Studi di Siena, Siena.

Macchi G., 2007, *Geografia dell'incastellamento*, All'Insegna del Giglio, Firenze.

Malanson G.P., 1999, *Considering complexity*, in «Annals of the Association of American Geographers», vol. 89, n. 4, pp. 746-753.

Mandelbrot B., 1967, *How long is the coast of Britain? Statistical self-similarity and fractional dimension*, «Science», vol. 156, n. 3775, pp. 636-637.

Mandelbrot B., 1982, *The fractal geometry of nature: updated and augmented*, Feeman, New York.

Matthews J.A., 1985, *Metodologia statistica per la ricerca geografica*, Franco Angeli, Milano.

McDermott I., O'Connor J., 2003, *Il pensiero sistemico*, Sperling & Kupfer, Milano.

Meadows P., 1957, *Models, systems and science*, «American Sociological Review», vol. 22, n. 1, pp. 3-9.

Merlin P., 1983, *Analisi quantitativa e spazio urbano*, Franco Angeli, Milano.

Micela G., Granata L., Iuliano V., 2001, *Due secoli di pioggia a Palermo*, Università di Palermo - Osservatorio Astronomico di Palermo Giuseppe S. Vaiana, Palermo.

Middleton R., Wardley P., 1990, *Information Technology in economic and social history: the computer as philosopher's stone or Pandora's box?*, «The Economic History Review», New Series, vol. 43, n. 4, pp. 667-696.

Miles R.E., 1972, *The random division of space*, «Advances in Applied Probability», vol. 4, pp. 243-266.

MITCHELL J., 1994, *Numbers as quantitative relations and the traditional theory of measurement*, «The British Journal for the Philosophy of Science», vol. 45, n. 2, pp. 389-406.

O'SULLIVAN D., UNWIN D.J., 2003, *Geographic information analysis*, Willey, Hoboken.

OLIVELLI F., 2006, *Insediamenti della Repubblica Senese: Tavola delle Possessioni, analisi spaziale e studi quantitativi*, Tesi di Laurea, Università degli Studi di Siena, Siena.

PARR J.B., 1973, *Structure and size in the urban system of Lösch*, «Economic Geography», vol. 49, n. 3, pp. 185-212.

PATTISON W.D., 1964, *The four traditions of geography*, «Journal of Geography», vol. 63, n. 5, pp. 211-216.

PASSERI V., NERI L., 1994, *Gli insediamenti della Repubblica di Siena nel catasto del 1318-1320*, Università degli studi di Siena - Dipartimento di Scienze Storiche, Giuridiche, Politiche e Sociali, Siena.

PEARSON C.E., 1980, *Rank size distributions and the analysis of prehistoric settlement systems*, «Jounal of Anthropological Research», vol. 36, n. 4, pp. 453-462.

POINCARÉ H., 1963, *Mathematics and science: last essays*, Dover Publications, New York.

POISSON S.D., 1837, *Recherches sur la probabilité des jugements en matières criminelles et matière civile*, Bachelier, Paris.

POLI D., 2001, *Attraversare le immagini del territorio: un percorso fra geografia e pianificazione*, All'Insegna del Giglio, Firenze.

POPPER K.R., 1994, *Poscritto alla logica della scoperta scientifica: il realismo e lo scopo della scienza*, Il Saggiatore, Milano.

POPPER K.R., 1995a, *Il mito della cornice*, Il Mulino, Bologna.

POPPER K.R., 1995b, *Modelli, strumenti e verità. Lo status del principio di razionalità nelle scienze sociali*, in POPPER K.R. 1995a, pp. 207-245.

POPPER K.R., 1996, *La conoscenza e il problema corpo-mente*, Il Mulino, Bologna.

POPPER K.R., 1997, *Miseria dello storicismo*, Feltrinelli, Milano.

POPPER K.R., 1998, *La logica della scoperta scientifica*, Einaudi, Torino.

Preston R.E., 1971, *The structure of central place systems*, «Economic geography», vol. 42, n. 2, pp. 136-155.

Prigogine I., 2006, *Le leggi del caos*, Laterza, Roma.

Quaini M., 1974, *Marxismo e geografia*, La Nuova Italia, Firenze.

Quaini M., 1975, *La costruzione della geografia umana*, La Nuova Italia, Firenze.

Rhind D., 1977, *Computer-aided cartography*, «Transactions of the Institute of British Geographers», New Series, vol. 2, n. 1, pp. 71-97.

Ripley B.D., 1976, *The second-order analysis of stationary point processes*, «Journal of applied probability», vol. 13, n. 2, pp. 255-266.

Ripley B.D., 1977, *Modelling spatial patterns*, «Journal of the Royal Statistical Society», Series B, vol. 39, n. 2, pp. 172-212.

Ripley B.D., 1979, *Test of "randomness" for spatial point patterns*, «Journal of the Royal Statistical Society», Series B, vol. 41, n. 3, pp. 368-374.

Ripley B.D., 1981, *Spatial statistics*, Wiley, New York.

Rodgers A., 1974, *The locational dynamics of soviet industry*, «Annals of the Association of American Geographers», vol. 64, n. 2, pp. 226-240.

Rombai L., 2002, *Geografia storica dell'Italia: ambienti, territori, paesaggi*, Le Monnier, Firenze.

Rushton G., 1969, *Analysis of spatial behavior by revealed space preference*, «Annals of the Association of American Geographers», vol. 59, n. 2., pp. 391-400.

Sack R.D., 1974, *Chorology and spatial analysis*, «Annals of the Association of American Geographers», vol. 64, n. 3., pp. 439-452.

Savage S.H., Falconer S., 2003, *Spatial and statistical inference of late bronze age polities in the southern Levant*, in «Bulletin of the American Schools of Oriental Research», n. 330, pp. 31-45.

Schaefer F.K., 1953, *Exceptionalism in geography: a methodological examination*, «Annals of the Association of American Geographers», vol. 43, n. 3, pp. 226-249.

Sharov A., 1992, *Life system approach: a system paradigm in population ecology*, «Oikos», LXIII, pp. 485-494.

SERENI E., 1989, *Storia del paesaggio agrario italiano*, Laterza, Roma-Bari.

SETTIA A., 1982, *Pievi e cappelle nella dinamica del popolamento rurale*, in *Cristianizzazione ed organizzazione ecclesiastica nelle campagne dell'Alto Medioevo: espansione e resistenze*, Atti delle XXVIII Settimane di studio del Centro Italiano di Studi sull'Alto Medioevo (Spoleto 1980), Spoleto, pp. 445-493.

SIBLEY D., 1981, *The notion of order in spatial analysis*, «The Professional Geographer», vol. 33, n. 1, pp. 1-5.

SILK J., 1979, *Statistical concepts in geography*, George Allen & Unwin, Londra.

SPIEGEL M.R., 1994, *Statistica*, McGraw-Hill, Milano.

SPIEKERMANN K., WEGENER M., 2000, *Freedom from the tyranny of zones: towards new GIS based spatial models,* in FOTHERINGHAM, WEGENER 2000, pp. 45-61.

STODDART D.R., 1986, *On geography*, Basil Blackwell, Oxford.

STOLPER W.F., 1959, *August Lösch in memoriam*, in LÖSCH 1959, pp. vii-xi.

STOPANI R., 1995, *Guida ai percorsi della via Francigena*, Le Lettere, Firenze.

STOW D.A., 1993, *The role of geographic information systems for landscape ecological studies,* in HAINES-YOUNG, GREEN, COUSINS 1993, pp. 11-22.

STRAHLER A.H., 1950, *Equilibrium theory of erosional slopes approached by frequency distribution analysis*, «American Journal of Science», vol. 248, pp. 673-696 e pp. 800-814.

SVIATLOVSKY E.E., EELLS C.W., 1937,*The centrographical method and regional analysis*, «Geographical Review», vol. 27, n. 2, pp. 240-254.

TABACCO G., 1974, *Egemonie sociali e strutture del potere nel medioevo italiano*, Einaudi, Torino.

TAYLOR P.J., 1977, *Quantitative methods in geography: an introduction to spatial analysis*, Houghton Mifflin, Boston.

TARRANT J.R., 1968, *A Note concerning the definition of groups of settelments for a central place hierarchy*, «Economic Geography», XLIV, n. 2, pp. 144-151.

Thiessen A.H., 1911, *Precipitations averages for large areas*, «Monthly Weather Review», vol. 39, n. 7, pp. 1082-1084.

Thomas D.H., 1971, *On use of cumulative curves and numerical taxonomy*, «American Antiquity», XXXVI, n. 2, pp. 206-209.

Tim P., Unwin H., 1981, *The rank-size distribution of medieval english taxation hierarchies with particular reference to Nottinghamshire*, «The Professional Geographer», vol. 33, n. 3, pp. 350-360.

Tinacci Mossello M., 1990, *Geografia economica*, Il Mulino, Bologna.

Tobler W.R., 1962, *A classification of map projections*, «Annals of the Association of American Geographers», vol. 52, n. 2, pp. 167-175.

Tobler W.R., 1970, *A computer movie simulating urban growth in the Detroit region*, «Economic Geography», vol. 46, n. 2, pp. 234-240.

Tolman C.P., 1930, *Review: Alfred Weber's theory of the location of industries*, «The American Economic Review», vol. 20, n. 1, pp. 110-111.

Toomer G.J., 1998, *Ptolemy's Almagest*, University Press, Princeton.

Tsai V.J.D., 1993, *Delaunay triangulations in TIN creation: an overview and a linear-time algorithm*, «International Journal of Geographical Information Science», vol. 7, n. 6, pp. 501-524.

Turnbull D., 1996, *Cartography and science in early modern Europe: mapping the construction of knowledge spaces*, «Imago Mundi», vol. 48, pp. 5-24.

Unwin D., 1986, *Analisi spaziale: un'introduzione geografica*, Franco Angeli, Milano.

Vagaggini V. (a cura di), 1978, *Spazio geografico e spazio sociale*, Franco Angeli, Milano.

Vagaggini V., 1982, *Le nuove geografie. Logica teoria e metodi della geografia contemporanea*, Herodote, Genova.

Vagaggini V., Dematteis G., 1976, *I metodi analitici della geografia*, La Nuova Italia, Firenze.

Vallega A., 1999, *Introduzione alla geografia umana*, Mursia, Milano.

Vallega A., 2004, *Geografia umana*, Le Monnier, Firenze.

WAGSTAFF J.M., 1978, *A possible interpretation of settlement pattern evolution in terms of "catastrophe theory"*, «Transactions of the Institute of British Geographers», New Series, vol. 3, n. 2, pp. 165-178.

WALLÉN C.C., 1955, *Some characteristics of precipitation in Mexico*, «Geografiska Annaler», vol. 37, n. 1-2, pp. 51-85.

WEBER A., 1929, *Theory of the location of industries*, Chicago University Press, Chicago.

WEBER A., 1968, *Theory of the location of industries*, Chicago University Press, Chicago.

WEBER M., 1979, *La città*, Bompiani, Milano.

WEGENER M., 2000, *Spatial models and GIS*, in FOTHERINGHAM, WEGENER 2000, pp. 3-20.

WICKEN J.S., 1987, *Entropy and information: suggestions for common language*, «Philosophy of Science», vol. 54, n. 2., pp. 176-193.

WILSON A.G., 1972, *Theoretical geography: some speculations*, «Transactions of the Institute of British Geographers», n. 57, pp. 31-44.

WOODBURY R.B., 1960, *Nels C. Nelson and chronological archaeology*, «American Antiquity», XXV, n. 3, pp. 400-401.

WRIGHT D.J., GOODCHILD M.F., PROCTOR J.D., 1997, *GIS: tool or science? Demystifying the persistent ambiguity of GIS as "tool" versus "science"*, «Annals of the Association of American Geographers», vol. 87, n. 2, pp. 346-362.

WRIGLEY N., 1977, *Probability surface mapping: a new approach to trend surface mapping*, «Transactions of the Institute of British Geographers», vol. 2, n. 2, pp. 129-140.

YI-FU TUAN, 1978, *Spazio e luogo, una prospettiva umanistica*, in VAGAGGINI 1978, pp. 92-130.

YOUNG G.L., 1978, *Hierarchy and central place: some questions of more general theory*, «Geografiska Annaler», Series B, Human Geography, LX, n. 2, pp. 71-78.

YOUNG T.C., 1988, *Since Herodotus, has history been a valid concept?*, «American Antiquity», LIII, n. 1, pp. 7-12.

Elenco delle figure

1.1 La “posizione” nella ricerca geografica 26
1.2 Composizione e distribuzione 30
1.3 Il ruolo dello spazio nelle dinamiche di stanziamento 34
1.4 La spostamento degli oggetti nello spazio 40
1.5 La distribuzione spaziale dei fenomeni 43
1.6 Cartografia e lettura delle distribuzioni 44
1.7 Complessità dei rapporti topologici 48
1.8 Comparazione spaziale 50
1.9 Differenze tra caos e casualità 51
1.10 Il modello rurale di Hudson 52

2.1 La sfera armillare di Keplero 64
2.2 S. Münster, *La Cosmographie universelle*, 1552. 66
2.3 Immagini come modelli cartografici 68
2.4 Il modello vettoriale dei dati spaziali 70
2.5 Confronto tra modello vettoriale e modello *raster* 70

3.1 I poligoni di Thiessen 79
3.2 L’area di pertinenza 80
3.3 La vicinanza geografica 81
3.4 Rappresentazione delle trame di Delaunay. 82
3.5 Poligoni di Thiessen e confini amministrativi 83
3.6 Gli anelli di Von Thünen 86
3.7 La produzione di legname secondo Von Thünen 88
3.8 La formazione dei confini fra gli anelli di Von Thünen 90
3.9 Anelli intorno ad Addis Abeba 91
3.10 Modello di Weber: aree industriali vicine ad aree di estrazione . 93
3.11 Modello di Weber: aree industriali vicine ai mercati 94
3.12 Modello di Weber con punti di carico e scarico 95
3.13 Il modello di Weber sul piano. 97
3.14 Il popolamento come risorsa nel modello di Weber. 98

3.15 Le *iso-time* nel modello di Weber. 99
3.16 Le isodapane. 100
3.17 Esempio isodapane. 101
3.18 I territori complementari di Christaller 104
3.19 La struttura ad alveare della maglia dei territori complementari 105
3.20 I ranghi inferiori secondo Christaller 106
3.21 Il modello di Christaller 108
3.22 Le tre aree complementari minime secondo Lösch. 111
3.23 Il paesaggio löschiano. 113
3.24 Struttura del paesaggio löschiano 115
3.25 Regola rango dimensione e distribuzione della popolazione reale 116
3.26 Logaritmi della distribuzione di Zipf e della distribuzione della popolazione reale 117
3.27 Modello binario e delle città primate 121
3.28 Il valore Θ nel modello di Zipf 123

4.1 Le scale di misurazione 130
4.2 Esempi cartografici di scale di misurazione nominali, ordinali e a rapporti 134
4.3 Tipi di campionamento 141
4.4 Carta dei comuni della provincia di Siena. 145
4.5 Istogramma dei valori relativi alla superficie dei comuni della provincia di Siena. 149
4.6 Poligono della distribuzione di frequenza delle superfici dei comuni della provincia di Siena. 150
4.7 Istogrammi ad intervalli non regolari 151
4.8 Classificazione cartografica 153
4.9 Curva cumulativa 156
4.10 Misure di tendenza centrale e misure di dispersione 157
4.11 Griglia e media ponderata 158
4.12 Misure di tendenza centrale 161
4.13 Distribuzioni di frequenza e tendenza centrale 163
4.14 Quartili 165
4.15 Scarto quadratico medio 168
4.16 Altezza della specie umana e distribuzione normale 175
4.17 Distribuzione normale e precipitazioni 177
4.18 Processo casuale indipendente 178
4.19 Distribuzione di Poisson 179

5.1 Densità e struttura spaziale 185
5.2 La desnità come strumento di confronto 186

5.3 Carta della densità abitativa dei comuni della provincia di Siena nel 1951 . . . 189
5.4 Carta della densità abitativa dei comuni della provincia di Siena nel 1961. . . . 191
5.5 Griglie di densità . . . 194
5.6 Suddivisione arbitraria tramite anelli concentrici . . . 201

6.1 Carta di distribuzione a punti . . . 205
6.2 Baricentro . . . 209
6.3 Carta della serie temporale del baricentro . . . 212
6.4 Scala di dispersione . . . 214
6.5 Struttura spaziale e scala geografica . . . 216
6.6 Misurazione di frequenza basate su griglie . . . 219
6.7 Scala di dispersione basata sull'analisi a quadrati . . . 220
6.8 Analisi a quadrati con *kernel* casuale . . . 223
6.9 Teorema di Pitagora . . . 225
6.10 Scala dell'indice R . . . 226
6.11 Rappresentazione cartografica dei vicini più prossimi . . . 228
6.12 Istogramma dei vicini prossimi . . . 229
6.13 Significatività dell'indice di dispersione R . . . 232
6.14 Limiti del metodo del vicino più prossimo . . . 233
6.15 Confini e funzione K . . . 237
6.16 La funzione K di Ripley . . . 238
6.17 La misurazione di frequenze con la funzione K di Ripley . . . 240
6.18 Funzione K e lo studio dei *patterns* rurali . . . 242

Realizzazione editoriale: Betti s.r.l.
Siena, Italia - Tel & fax 0577 289447
www.betti.it - info@betti.it

Finito di stampare nel mese di marzo 2009.

Uno degli elementi più importanti di qualsiasi processo scientifico è lo scambio di idee e osservazioni sulla realtà all'interno di una cornice di riferimento chiara e precisa. A livello epistemologico, la costruzione di «nuova conoscenza» si fonda sulla capacità di una disciplina di comunicare concetti e teorie in modo semplice, al di là di possibili interpretazioni soggettive. Il sapere tende a progredire in contesti caratterizzati da sistemi di trasmissione obiettivi, ma, soprattutto, condivisi.
Nel sottolineare il ruolo dei linguaggi formali come strumento essenziale per la comunicazione di idee e teorie relative allo spazio, *Spazio e misura* rappresenta una guida per coloro che intendono avvicinarsi all'applicazione dei metodi quantitativi in geografia. Mettendo l'accento proprio sul problema dei linguaggi, il testo illustra un percorso introduttivo chiaro e lineare, pensato per studenti dell'area umanistica. Il volume, infatti, nel rendere semplice il processo di lettura e approfondimento, cerca di spiegare passo per passo i singoli dettagli metodologici, al di fuori di qualsiasi tecnicismo statistico-matematico, e di illustrarne, con una serie di esempi concreti, le diverse applicazioni in ambito geografico.

Giancarlo Macchi Jánica insegna geografia presso la Facoltà di Lettere e Filosofia dell'Università degli Studi di Siena. Si occupa dello studio delle reti dell'insediamento rurale e della comprensione dell'evoluzione spaziale del popolamento. Da più di un decennio lavora nell'ambito dei metodi quantitativi applicati alla geografia umana e si dedica all'applicazione dei sistemi GIS allo studio della storia dell'insediamento.
È autore di *Geografia dell'incastellamento* (2007) oltre che di diversi saggi tra i quali *La struttura della maglia dei castelli medievali nell'Italia centrale: paralleli tra modelli di stanziamento umano* (2006)*; Modelli, modelli geografici e modelli scientifici* (2007)*; Introduzione all'analisi spaziale* (2003)*; Sulla misurazione delle forme d'occupazione sociale dello spazio medievale* (2001).

Edizioni dell'Università - Siena
ISBN 978-88-96151-01-3